Review of the Development Effectiveness of the United Nations Development Programme (UNDP)

2005 - 2011

13 April 2012

Preface

This review provides an independent and evidence-based assessment of the development effectiveness of the United Nations Development Programme (UNDP). The approach to reviewing multilateral organizations was developed under the guidance of the Organisation for Economic Co-operation and Development (OECD)'s Development Assistance Committee Network on Development Evaluation (DAC-EVALNET). It complements the periodic assessments done by the Multilateral Organization Performance Assessment Network of the multilateral organizations' operational effectiveness.

The review was led by the Evaluation Directorate of the Canadian International Development Agency (CIDA), with support from the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs. The methodology applied in the review is a structured meta-synthesis of the content of 55 evaluations published by UNDP's Evaluation Office using a common set of assessment criteria derived from the DAC evaluation criteria. It was conducted by a team from CIDA's Evaluation Directorate (Colum Grove-White, Tricia Vanderkooy, and Andres Velez-Guerra) and the consulting firm, Goss Gilroy Inc. (Chris Khng, Ted Freeman, Danielle Hoegy, and Louise Mailloux). The team is grateful to UNDP for its helpfulness and useful, practical advice.

This report is designed for use by all members of the DAC-EVALNET and other interested stakeholders. The findings of this review are also presented in a separate report prepared for CIDA, which includes information on CIDA's relationship with UNDP. A similar review of the development effectiveness of the World Food Programme has also been conducted and we hope that these two reviews represent the beginning of a series of reviews of multilateral organization effectiveness, based on their own evaluations.

We would like to thank everybody who contributed to this review for their valued input and support.

Ted Kliest
Senior Evaluator
Policy and Operations Evaluation Department
Netherlands Ministry of Foreign Affairs

Goberdhan Singh
Director General
Evaluation Directorate
Canadian International Development Agency

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADR   Assessment of Development Results

CIDA   Canadian International Development Agency

DAC/EVALNET   Development Assistance Committee Network on Development Evaluation

MDG   Millennium Development Goals

MOPAN   Multilateral Organization Performance Assessment Network

OECD/DAC   Development Assistance Committee of the Organization for Economic
                Development and Cooperation

UN   United Nations

UNDP   United Nations Development Programme

Executive Summary

Background

This report presents the results of a review of the development effectiveness of the United Nations Development Programme (UNDP). UNDP is the principal development agency of the United Nations (UN). UNDP is a solutions-oriented, knowledge-based development organization that supports countries in reaching their own development objectives and internationally agreed-upon goals, including the Millennium Development Goals (MDGs).[1] UNDP is a complex organization with operations in 177 countries and annual program expenditures of $4.4 billion USD. UNDP's strategic plan for 2008-2013 and its most recent annual report emphasize four pillars of its organizational mandate: poverty reduction and achievement of the MDGs; democratic governance; crisis prevention and recovery; and, environment and sustainable development. In addition to work on these four pillars, UNDP's work on promoting gender equality, women's empowerment, and South-South cooperation also supports partner countries in strengthening their own capacities.[2]

Purpose and Objectives

The purpose of the review is to provide an independent and evidence-based assessment of the development effectiveness of UNDP programs for use by external stakeholders, primarily bilateral development agencies.

Methodology

The approach and methodology used in this review were developed under the guidance of the Development Assistance Committee's Network on Development Evaluation (DAC-EVALNET). It was developed to fill an information gap regarding the development effectiveness of multilateral organizations. Although these organizations produce annual reports to their management and/or boards, member states were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. This approach complements the organizational effectiveness assessment of the Multilateral Organization Performance Assessment Network (MOPAN).

The methodology does not rely on a particular definition of development effectiveness. The Management Group and the Task Team that developed the methodology had previously considered whether an explicit definition of development effectiveness was needed. In the absence of an agreed upon definition of development effectiveness, the methodology focuses on some of the essential characteristics of developmentally effective multilateral programming, as described below:

- Programming activities and outputs would be relevant to the needs of the target group and its members;
- Programming would contribute to the achievement of development objectives and expected development results at the national and local level in developing countries (including positive impacts for target group members);

---

[1] UNDP Annual Report – 2010-2011: People-centred Development, UNDP, 2011
[2] UNDP Strategic Plan – 2008-2011: Accelerating Global Progress on Human Development, UNDP. 2008

- Benefits experienced by target group members and the development (and humanitarian) results achieved would be sustainable in the future;
- Programming would be delivered in a cost efficient manner;
- Programming would be inclusive by supporting gender equality and being environmentally sustainable (thereby not compromising the development prospects in the future); and,
- Programming would enable effective development by allowing participating and supporting organizations to learn from experience and use tools such as evaluation and monitoring to improve effectiveness over time.

Based on the above-mentioned characteristics, the review's methodology uses a common set of assessment criteria derived from the DAC's existing evaluation criteria (Annex 1). It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organization. In early 2011, the methodology was endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations.

The core methodology used in the review is a structured meta-synthesis[3] of the content of a representative sample of 55 evaluations published by UNDP between 2009 and 2011 (Annex 2). The sample was drawn from a universe of 199 evaluations published by UNDP during the same period.[4] This universe was chosen to allow for maximum coverage of programming during the current UNDP strategic plan (2008-2013). In fact, the majority of the reviewed evaluations (67%) concentrated on the program period from 2005 to 2011.

Most of the evaluations in the sample (37 of 55) were either country program evaluations or global thematic evaluations carried out by UNDP's Evaluation Office (i.e. centralized evaluations). Of the remaining 18 evaluations, 15 were decentralized evaluations carried out by the UNDP office responsible for the program being evaluated. Two evaluations were led by the country concerned (Tanzania and South Korea respectively) and one was a joint evaluation managed collectively by UNDP, UNAIDS and the World Bank. The sampling process is described in further detail in the methodology annex (Annex 3).

After being screened for quality (the meta-evaluation[5] component of the review – see Annex 4), each evaluation was reviewed to identify findings relating to six main criteria for assessing the development effectiveness of UNDP-supported programs at the field level, namely:
- The Achievement of Development Objectives and Expected Results;
- Cross Cutting Themes: Inclusive Development which is Sustainable;
- Sustainability of Results/Benefits;
- Relevance of Interventions;
- Efficiency; and,
- Using Evaluation and Monitoring to Improve Development Effectiveness.

The criteria were assessed using 19 sub-criteria that are considered essential elements of effective development. Findings for each of the 19 sub-criteria were classified using a four-point scale: "highly satisfactory," "satisfactory," "unsatisfactory" and "highly unsatisfactory." Classification of findings was guided by a grid with specific instructions for each rating across all sub-criteria (Annex 5).

---

[3] "Meta-synthesis" refers to a compilation of findings from evaluations that have been vetted for quality.
[4] Excluding project evaluations.
[5] "Meta-evaluation" refers to the evaluation of the evaluations, or the quality control process conducted on the evaluations in the sample (Further details are provided in Annex 4).

Key Findings

Evaluations report that most UNDP programs achieve their development objectives and expected results

The reviewed evaluations report that most UNDP programs are achieving their development objectives and expected results, although weaknesses remain in some areas. Two thirds of evaluations reported findings of "satisfactory" or "highly satisfactory" for the achievement of development objectives and expected results (66%) and for reaching a substantial number of beneficiaries (65%). Sub-criteria on the positive benefits for target group members and significant changes in national development plans and policies were more highly rated (83% and 71%).

The most frequently cited factor hindering objectives achievement was dispersion of UNDP country programming across too many projects, too wide a geographic area or too many institutions. Factors contributing to UNDP effectiveness in achieving development objectives include effective investment in knowledge development; consultation to strengthen support for priority policies; and effective advocacy for the MDGs.

UNDP works to mainstream gender equality and promotes environmental sustainability

The result for UNDP's effectiveness in supporting gender equality is generally positive, with 62% of the evaluations reporting that programs performed to a "satisfactory" or "highly satisfactory" level. Evaluations also reported that UNDP programs have made progress on mainstreaming gender equality into program components and increasing women's participation in governance structures. An important factor hindering greater program success in gender equality was the absence of a systematic approach to gender analysis during the design and implementation of some programs. Factors contributing to success in the area of gender equality include a strategy of combining women-specific programs at a local or national level with mainstreaming of gender equality into all UNDP programs.

UNDP's effectiveness in supporting environmentally sustainable development was also rated positively. It is the third highest ranked sub-criteria in terms of percentage of evaluations reporting "satisfactory" or "highly satisfactory" findings (79%). Factors contributing to success in this area include UNDP's effective advocacy for environmental issues and its willingness to invest in environmental research methods and tools.

The sustainability of benefits is a significant challenge

The findings on sustainability of benefits/results represent a significant challenge to the development effectiveness of UNDP. In particular, the results for the likelihood of benefits continuing after program completion are negative (only 36% of evaluations reporting "satisfactory" or "highly satisfactory" results). This was the fourth lowest ranked sub-criteria measuring development effectiveness.

The evaluation findings on institutional and community aspects of sustainability are more positive but still indicate significant weaknesses in UNDP's effectiveness in supporting institutional and community capacity for sustainability, with one-fifth of evaluations reporting "highly unsatisfactory" results. However, UNDP programs have made a positive contribution to the enabling environment for development (75% of evaluations reporting "satisfactory" or "highly satisfactory" results). In particular, UNDP contributed to the enabling environment for development through its positive influence on national development planning, budgeting and programming.

A factor contributing to success in the sustainability of benefits/results was strong program ownership by the developing country partners. However, factors hindering sustainability were cited more frequently, including the absence in some programs of an explicit program phase-out strategy and the lack of integrated sustainability considerations in the early stages of program design (i.e. quality at entry[6] issues).

UNDP programs are relevant to the context in which they work

The three sub-criteria measuring relevance of interventions were rated "satisfactory" or "highly satisfactory" more frequently than those for any other main criteria. Most evaluations (73%) reported positive findings on the suitability of programs to the needs of target group members and 77% reported positive findings on effective partnerships with government, civil society and development partners. The alignment of UNDP programs with national development goals was ranked highest among all of the 19 sub-criteria (88%).

One factor contributing to success in ensuring the relevance of interventions was the use of consultation among stakeholders to build a consensus on specific needs and solutions. Nonetheless, improvements could be made by developing a more systematic approach to needs and risk assessments.

Evaluations report weaknesses in program efficiency

In comparison with results in other areas, the reviewed evaluations were negative in their assessment of UNDP's efficiency. The cost efficiency of UNDP program activities was rated "satisfactory" or better in just 52% of the evaluations reviewed. Similarly, UNDP timeliness was rated "satisfactory" or better in 51% of the evaluations reviewed. The efficiency of systems for program implementation was the third lowest ranked of all sub-criteria in terms of the percentage of evaluations reporting findings of "satisfactory" or better (27%).

The most commonly cited factor that hindered success in the area of efficiency was reportedly complex, rigid and overly bureaucratic UNDP systems, processes and procedures for project administration and control, particularly in relation to procurement and the disbursement of funds.

UNDP faces challenges in strengthening decentralized systems for evaluation, monitoring and results-based management

The review indicates that a significant ongoing challenge for UNDP is its weakness in evaluation systems, results monitoring, and results-based management at the decentralized program level. Decentralized systems and processes for evaluation were rated "satisfactory" or "highly satisfactory" in only 41% of the evaluations reviewed. Findings were even more negative for monitoring systems and the effectiveness of results-based management systems. These two sub-criteria were ranked lowest and second lowest of all sub-criteria (23% and 24%, respectively).

The most frequently cited factors hindering success included a tendency to focus results models, indicators and reporting systems on program activities and outputs rather than on outcomes, and the absence of either evaluation studies or evidence-based monitoring reports at the level of national and regional programs and projects to inform decision making.

However, UNDP does make systematic use of evaluation findings to improve development effectiveness, with many evaluations receiving a detailed management response (81% rated "satisfactory" or higher). This does not mean that poorly conducted or weak decentralized

---

[6] "Quality at entry" refers to front-end analysis, risk assessment, and definition of expected results.

evaluations are used to inform development programming. UNDP's Evaluation Office commissions, carries out and follows up on the quality of evaluations, which provide a reasonably strong base of evaluation reports being used to inform decision-making at the organizational level.

Conclusions: The Development Effectiveness of UNDP

The assessment of the development effectiveness of UNDP has demonstrated that most UNDP programs are achieving their development objectives and expected results, although weaknesses remain in some areas. Generally positive results were achieved when UNDP program performance was measured against the criteria for cross cutting themes of gender equality (62% "satisfactory" or better) and promoting environmental sustainability (79%). UNDP programs have also demonstrated relevance to the needs of target group members (73%), alignment with national development priorities (88%), and effectiveness in working with government, civil society and development partners (77%).

However, with one third of evaluations reporting objectives achievement as "unsatisfactory" or "highly unsatisfactory," UNDP faces challenges in improving the development effectiveness of its programming. Achieving development objectives was sometimes hindered by dispersion of UNDP programming across too many projects, too wide a geographic area or too many institutions. Gender analysis was also inconsistently applied in program design, limiting UNDP's effectiveness in gender equality.

UNDP also faces a significant challenge to development effectiveness in the limited sustainability of program results and benefits, as only 36% of evaluations reported "satisfactory" or "highly satisfactory" performance in this area. The lack of explicit program phase-out strategies and integration of sustainability into the early stages of program design hindered success in this area. Results on efficiency also indicate unsatisfactory performance, often attributed to complex, rigid and overly bureaucratic UNDP systems. Evaluation, monitoring and results-based management systems at the decentralized level were assessed as "unsatisfactory" or "highly unsatisfactory" in many of the reviewed evaluations. Nonetheless, UNDP's Evaluation Office produces quality evaluation reports that are being used to strengthen UNDP policies and programs. A survey of changes initiated by UNDP in the recent past is outside the scope of this review. However, UNDP provided an overview of the most important of these recent changes (Annex 7).

Areas for Attention

This analysis of the strengths and weaknesses of the UNDP's development programming led to the identification of areas that required ongoing attention from UNDP, in collaboration with its bilateral partners, including:

- Achieving greater focus in country programs by concentrating on fewer projects within a given country;
- Improving systems and processes for gender analysis in order to more effectively mainstream gender equality into development programs;
- Integrating project phase-out strategies and sustainability designs in the early staged of program development;
- Improve the flexibility and adaptability of UNDP systems for project and program administration and control; and,
- Further strengthening decentralized systems for evaluation, monitoring and results-based-management.

1.0 Background

1.1 Introduction

This report presents the results of a review of the development effectiveness of the United Nations Development Programme (UNDP). The approach and methodology were developed under the guidance of the DAC Network on Development Evaluation (DAC-EVALNET). It relies on the content of published evaluation reports produced by UNDP to assess development effectiveness.

The methodology uses a common set of assessment criteria derived from the DAC's existing evaluation criteria (Annex 1). It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organization. In early 2011, the overall approach and methodology were endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations. This review is among the first to be implemented following the DAC-EVALNET's endorsement.

From its beginnings, the process of developing and implementing the reviews of development effectiveness has been coordinated with the work of the Multilateral Organization Performance Assessment Network (MOPAN). By focusing on development effectiveness and carefully selecting assessment criteria, the reviews seek to avoid duplication or overlap with the MOPAN process. Normal practice has been to conduct the development assessment review in the same year as a MOPAN survey for any given multilateral organization. In the case of UNDP, a MOPAN survey was last completed in 2009 and another is currently scheduled for 2012.

1.2 Why conduct this review?

The purpose of the review is to provide an independent, evidence-based assessment of the humanitarian and development effectiveness of WFP operations (hereafter referred to as "programs") for use by external stakeholders, primarily bilateral development agencies.

The current approach to assessing the development effectiveness of multilateral organizations was developed in order to address a gap in the information available to bilateral development agencies. While MOPAN provides regular reports on the organizational effectiveness of multilateral organizations, it does not yet report on development effectiveness and has not addressed the information gap this review is meant to fill. Other options such as large-scale, joint donor-funded evaluations of a given multilateral organization are much more time-consuming, more costly and result in a significant management burden to the organization being evaluated before, during and after the evaluations. The current approach is intended to work in a coordinated way with initiatives such as the DAC-EVALNET/United Nations Evaluation Group (UNEG) Peer Reviews of multilateral organization evaluation functions. It also recognizes that multilateral organizations continue to make improvements and strengthen their reporting on development effectiveness. The ultimate aim of the approach is to be replaced by regular, evidence-based, field-tested reporting on development effectiveness provided by multilateral organizations themselves.

1.3   UNDP: A global organization committed to development

1.3.1 Background

As of 2011, UNDP is the largest UN agency, with 129 Country Offices and operations in 177 countries. It was established as the UN's principal development agency in 1966. According to its most recent annual report,[7] UNDP is a solutions-oriented, knowledge-based development organization that supports countries in reaching their own development objectives and internationally agreed-upon goals, including the Millennium Development Goals (MDGs).

1.3.2 Strategic Plan

UNDP's strategic plan for 2008-2013 and its most recent annual report emphasize four pillars of its organizational mandate:

- poverty reduction and achievement of the MDGs;
- democratic governance;
- crisis prevention and recovery; and,
- environment and sustainable development.

In addition to work on these four pillars, UNDP promotes gender equality, women's empowerment, and South-South cooperation, while also supporting partner countries in strengthening their own capacities in these areas.[8]

1.3.3 Description of UNDP's work, geographic coverage and funding

UNDP has one of the widest footprints of any multilateral organization, with 129 country offices and operations in 177 countries. The scale of UNDP's work is illustrated by its substantial expenditures across five different program areas[9] in 2010 (Table 1).[10] In 2010, UNDP's total program expenditures in developing countries accounted for $4.4 billion USD.

---

[7] UNDP Annual Report – 2010-2011: People-centred development, UNDP, 2011.

[8] UNDP Strategic Plan – 2008-2011: Accelerating Global Progress on Human Development, UNDP. 2008. This plan was later extended until 2013.

[9] The expenditures classified under "mitigating the impact of HIV/AIDS" can be considered part of the first pillar of UNDP's mandate (poverty reduction and promotion of the MDGs). Program areas are reported here according to UNDP's own expenditure reporting categories.

[10] Midterm review of the UNDP strategic plan and annual report of the administrator, UNDP, 2011.

Table 1: UNDP Expenditures in Five Program Areas (2010)

Program Area	Number of Countries	2010 Expenditures (Millions USD $)
Poverty Reduction and MDG Achievement	135	1,349
Mitigating the Impact of AIDS	56	349
Democratic Governance	128	1,184
Crisis Prevention and Recovery	86	1,052
Environment and Sustainable Development	124	508
Total	-	**4,442**

UNDP funding contributions are provided from four distinct sources[11]. The four types of funding are: a) regular contributions; b) co-financing through cost-sharing and trust fund contributions by bilateral donors; c) contributions from multilateral partners; and, d) co-financing contributions by program countries.

For the three year period 2008-2010, UNDP reported total contributions from all sources of 14.4 billion USD. As illustrated by Figure 1, only 21% of the funding to UNDP during the period came in the form of regular contributions. Taken together, trust funds and co-financing by bilateral donors and co-financing by program countries accounted for 49% of UNDP funding during this three year span.

Figure 1: UNDP Financing by Source, 2008 to 2010 (in percentages)[12]

---
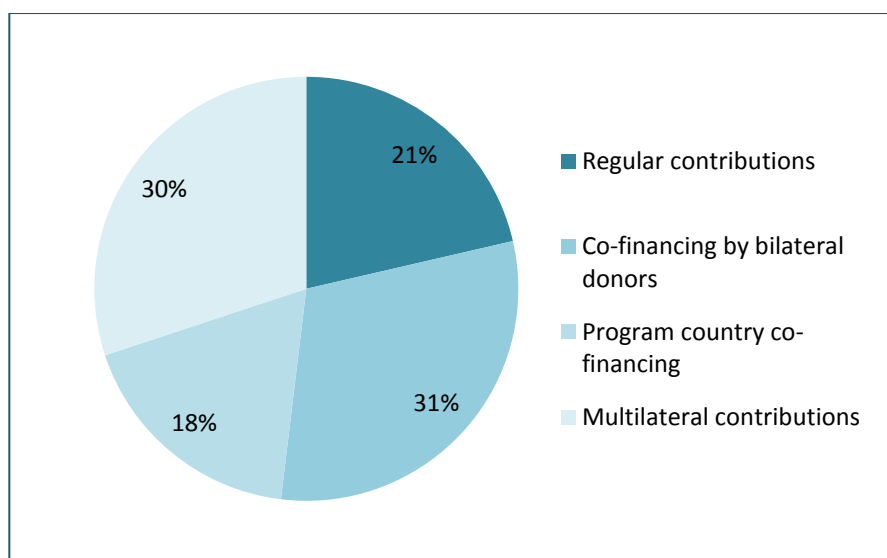
[11] Midterm review of the UNDP strategic plan and annual report of the administrator. UNDP, 2011. p.38.
[12] Midterm review of the UNDP strategic plan and annual report of the administrator. UNDP, 2011.

1.3.4 Evaluation and Results Reporting

UNDP's Evaluation Office

UNDP's Evaluation Office is responsible for conducting independent evaluations of corporate and global, regional, and country program outcomes identified in the UNDP strategic plan. In 2010/11, the office conducted 14 Assessments of Development Results (ADRs). On the other hand, outcome oriented decentralized evaluations are carried out by the program responsible, including regional and country offices where appropriate.

In addition, under the recently revised policy on evaluation at UNDP, the Evaluation Office acts as the corporate custodian of the evaluation function and its governance and accountability. Among other roles, it reports annually to the executive board on the function, compliance, coverage and quality of evaluations conducted by UNDP. It also maintains a system to make all evaluation reports, management responses and the status of follow up actions publicly accessible.

The Evaluation Office also works to "build a culture of evaluation in UNDP."[13] In doing so, it develops training materials for use by all organizational units responsible for evaluation plans. In 2011, the Evaluation Office began assessing the quality of all evaluation reports commissioned in accordance with respective evaluation plans in the organization. It also maintains the online roster of evaluation experts at UNDP and manages the organization's knowledge network on evaluation. All of these activities have the effect of providing support to decentralized evaluations, which is an acknowledged challenge for UNDP in the 2010 Report on Evaluation, as submitted to the Executive Board in June 2011.

UNDP's Evaluation Office implements the following types of independent evaluations:

- *Thematic Evaluations* assess UNDP performance in areas that are critical to contributing to development results at the global and regional level, including policies, focus areas, partnerships, program approaches, cooperation modalities or business models.
- *Global, regional and South-South program evaluations* assess the performance and intended results of these programs.
- *Assessments of Development Results (ADRs)* assess the attainment of intended and achieved results as well as UNDP contributions to development results at the country level.
- *Joint Evaluations* are a modality for carrying out an evaluation to which different development partners contribute. Any evaluation can be conducted as a joint evaluation.

In addition to these centralized evaluations, UNDP's decentralized regional and country-level program units conduct the following decentralized evaluations:

- *Thematic Evaluations* assess UNDP's performance in areas that are critical to contributing to development results in a given context that is pertinent to a UNDP program unit. They may cover UNDP's results and focus areas, crosscutting issues, partnerships, program approaches, cooperation modalities or business models.

---

[13] *Annual Report on Evaluation* in UNDP: 2010. UNDP, 2011. p.3.

- *Global, regional and country program evaluations* assess UNDP's attainment of intended results and contributions to development results.
- *UNDAF Evaluations* focus on United Nations Development Assistance Framework (UNDAF) outcomes, their contribution to national priorities and the coherence of United Nations Country Team support.
- *Outcome evaluations* address the short term, medium term and long-term results of a program or cluster of related UNDP projects.
- *Project evaluations* assess the efficiency and effectiveness of a project in achieving its intended results.

A Peer Review of the Evaluation Function at UNDP was carried out in 2005 under the auspices of the DAC Evaluation Network/UN Evaluation Group. The review addressed the question: "Does UNDP's Evaluation Office produce evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognized evaluation peers?" The peer review's authors responded that:

> "The United Nations Development Programme has an Evaluation Office 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." [14]

Reporting on Development Effectiveness

UNDP reports to its Executive Board on development effectiveness mainly through three regular agency-wide reporting documents: the UNDP annual report, the *Annual Report on Evaluation* (ARE), and the Administrator's regular report on the progress of the strategic plan.

The UNDP annual report[15] highlights global activities and provides a qualitative picture of development results. This picture is further refined through detailed case examples of work in specific countries on the four strategic program areas of UNDP's mandate. Quantitative information is provided throughout the report, mainly at the output level. The report describes the extensive reach of UNDP programming as well as the number and type of beneficiaries reached. It also provides information on the portions of UNDP funding and programming going to each of the four priority program areas and to crosscutting themes such as gender equality.

The *Annual Report on Evaluation* provides clear information on the annual level of evaluation activity in all different program areas of UNDP, as well as an assessment of the quality of evaluation reports. The *Annual Report on Evaluation* clearly identifies the roles of the Evaluation Office and other offices responsible for the evaluation function. In addition, the *Annual Report on Evaluation* provides an overview of the findings of evaluations carried out each year and identifies the challenges faced by UNDP in improving its efficiency and effectiveness.

---

[14] Peer Assessment of Evaluation in Multilateral Organizations: United Nations Development Programme. Ministry of Foreign Affairs of Denmark. 2005, p. 4.
[15] People Centered Development: UNDP in Action - Annual Report 2010/2011, UNDP, 2011.

The 2010 report noted positive findings around UNDP's role as a trusted and valued partner contributing to human development in the program countries. It also noted, however, challenges in the following areas:

- Making better use of "downstream" project work by better considering and planning for opportunities to "scale up" to the national level;
- Finding the appropriate balance between short-term and long-term project investments in a given country;
- The need to better address gender issues by providing greater attention and resources and by using appropriate analysis to formulate concrete strategies;
- The need for better defined results frameworks and more concrete baselines and performance measures in project design;
- The need to develop better strategies for "graduation" and for the scaling up of pilot activities;
- The need for stronger monitoring and evaluation, knowledge management and guidance, especially for country offices; and,
- The challenge of business and operational practices, leading to delays in payments to partners or in procurement and subsequent delays in the timely completion of planned projects. This further compromises the overall efficiency and effectiveness of UNDP interventions.

Many of these challenges are also noted in the findings and conclusions of this report.

At a substantive level, the Administrator provides the Executive Board with regular reports on UNDP's progress in implementing the strategic plan. In 2011, there was a significant improvement in the evidence base used to report on the strategic plan. The *Midterm Review of the UNDP Strategic Plan and Annual Report of the Administrator*[16] was provided to the Executive Board in advance of its meeting in June 2011. This report makes use of the work of UNDP's Evaluation Office by making frequent references to the findings of ADRs, thematic, and outcome evaluations.

For each priority programming area, the *Midterm Review* draws on the results of a set of applicable evaluation reports (from a low of seven to a high of 13 evaluations depending on the program area). The information extracted from the evaluations includes a summary of findings for each outcome, along with examples of successes and challenges.

This approach to reporting on development effectiveness draws on UNDP's own body of published evaluative evidence as summarized in the ARE. It goes some way to meeting the need for information on development effectiveness that is backed by field-based evidence. Unfortunately, the number of evaluations used is small and findings in the *Annual Report on Evaluation* are not quantified. The *Midterm Review* does highlight some of the challenges facing UNDP in its pursuit of the goals of the strategic plan. However, it lacks a balanced reporting of the distribution of evaluation findings (both positive and negative) so that the reader could appreciate the significance of positive and negative findings.

---

[16] Midterm Review of the UNDP Strategic Plan and Annual Report of the Administrator, UNDP, 2011.

In summary, UNDP's own regular reporting on development effectiveness, although strengthened in 2011, currently lacks a balanced overview of field-tested evidence on development effectiveness. In addition, the present review reflects many of the challenges identified in the 2010 ARE.

2.0 Methodology

This section briefly describes key aspects of the review's methodology. A more detailed description of the methodology is available in Annex 3.

2.1 Rationale

The term "common approach" describes the use of a standard methodology, as implemented in this review, to consistently assess the development effectiveness of multilateral organizations. It offers a rapid and cost effective way to assess the development effectiveness of the organization, relative to a more time consuming and costly joint evaluation.[17] The approach was developed to fill an information gap regarding development effectiveness of multilateral organizations. Although these organizations produce annual reports to their management and/or boards, member states were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. This approach complements MOPAN's organizational effectiveness assessment of multilateral organizations. The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[18]

1. There is a need for field-tested and evidence-based information on the development effectiveness of the multilateral organization.

2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-synthesis[19] methodology to synthesize an assessment of the organization's development effectiveness.

The first condition is satisfied, as UNDP's existing reporting mechanisms do not provide sufficient information on the organization's development effectiveness. The second condition is also satisfied, as the Evaluation Office at UNDP does produce enough evaluation reports of good quality and with sufficient coverage of investments to support an assessment of the development effectiveness of UNDP. Further details on the rationale are available in Annex 3.

---

[17] "Joint evaluation" refers to a jointly funded and managed comprehensive institutional evaluation of an organization. It does not refer to DAC/UNEG Peer Reviews of the Evaluation Function.

[18] Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

[19] "Meta-synthesis" refers to a compilation of findings from evaluations that have been vetted for quality.

2.2 Scope

The review is based on a sample of 55 evaluations selected from those published by UNDP from the beginning of 2009 to August 2011. The sample was selected from a universe of 199 higher level (non-project) evaluations reports published by UNDP in the same period[20]. Most of the evaluations in the sample (37 of 55) were either Assessment of Development Results (ADRs) or global thematic evaluations carried out by UNDP's Evaluation Office (i.e. centralized evaluations). Of the remaining 18 evaluations, 15 were decentralized evaluations carried out by the UNDP office responsible for the program being evaluated. Two evaluations were led by the country concerned (Tanzania and South Korea respectively) and one was a joint evaluation managed collectively by UNDP, UNAIDS and the World Bank. See Annex 3 for further details on sampling.

The 2009 to 2011 period was chosen to allow for maximum coverage of UNDP operations during the current strategic plan (2008 to 2013). The reviewed evaluations cover a lengthy programming period (i.e. from 2000 to 2011). However, the majority (70%) of the reviewed evaluations concentrated on the program period from 2005 to 2011. Therefore, the review team is confident that the evaluations reviewed report findings relevant to the current UNDP strategic plan. See Annex 3 for further details on the time period covered by the reviewed evaluations.

The evaluations reviewed cover all four priority areas of UNDP programming (poverty reduction, democratic governance, crisis prevention and recovery, and environment and sustainable development). The evaluation sample included 31 country program evaluations (the ADRs). Together, these 31 countries accounted for 1.2 billion USD of expenditures in 2009 or 33% of UNDP's total expenditures in the same year. In addition, the sample of evaluations included thematic evaluations that were regional or global in scope.

The review of evaluation reports was supplemented by a review of UNDP corporate documents. Information gathered through the review of UNDP documents was used to contextualize the results of the review. A list of the documents consulted is provided in Annex 6.

2.3 Criteria

The methodology involves a systematic and structured meta-synthesis of the findings of a sample of UNDP evaluations, according to six main criteria and 19 sub-criteria that are considered essential elements of effective development (Annex 1). The main criteria and sub-criteria are derived from the DAC Evaluation Criteria, with further refinements made during the pilot testing of the Asian Development Bank and the World Health Organization.

The methodology does not rely on a particular definition of development effectiveness. The Management Group and the Task Team that developed the methodology had previously considered whether an explicit definition of development effectiveness was needed. In the absence of an agreed upon definition of development effectiveness, the methodology focuses

---

[20] As of February 2012, UNDP's Evaluation Resource Centre indicates that 822 evaluations were electronically published in the 2009 to 2011 period, out of which 784 concerned UNDP programs (with the remainder focused on United Nations Volunteers or United Nations Capital Development Fund operations). The majority of these UNDP evaluations were project evaluations, which were not considered for inclusion in the sample of evaluations used by the Review.

on some of essential characteristics of developmentally effective multilateral organization programming, as described below.

- – the programming activities and outputs would be relevant to the needs of the target group and its members;
- – the programming would contribute to the achievement of development objectives and expected development results at the national and local level in developing countries (including positive impacts for target group members);
- – the benefits experienced by target group members and the development (and humanitarian) results achieved would be sustainable in the future;
- – programming would be delivered in a cost efficient manner.
- – the programming would be inclusive by supporting gender equality and being environmentally sustainable (thereby not compromising the development prospects in the future); and,
- – the programming would enable effective development by allowing participating and supporting organizations to learn from experience and use tools such as evaluation and monitoring to improve effectiveness over time.

Criteria for Assessing the Development Effectiveness of UNDP

- • The Achievement of Development Objectives and Expected Results;

- • Cross Cutting Themes (Environmental Sustainability and Gender Equality);

- • Sustainability of Results/Benefits;

- • Relevance of Interventions;

- • Efficiency; and

- • Using Evaluation and Monitoring to Improve Development Effectiveness.

2.4   Limitations

As with any meta-synthesis, major concerns include sampling bias, ensuring adequate coverage of the criteria, and classification problems. Sampling bias was minimized by providing adequate coverage of UNDP's published evaluations, including evaluations of national, regional and global programs. The 19 sub-criteria are well covered in the evaluations reviewed (See Annex 4 for further information).

First, a limitation arises from the classification of evaluation findings. ADRs often report their findings according to the main UNDP programming areas in a given country. Thus, arriving at a summary finding for the overall country program based on an ADR evaluation usually requires combining results from across various program areas, since results often differ by program area. ADR reports often merged results across the main program areas into a single summative finding although this was sometimes not made explicit in the evaluation report. On

the few occasions when an evaluation report did not include a clear summative finding on, for example, objectives achievement, the review team compiled evidence and made an overall assessment themselves.

Second, a limitation of the review is weak coverage of the criteria assessing efficiency. Several evaluations did not report on two of the efficiency sub-criteria, resulting in a moderate level of coverage. For sub-criteria 5.1, regarding the cost efficiency of program activities, 39 evaluations addressed this issue, and for sub-criteria 5.2 on the timeliness of program implementation, 31 evaluations addressed this issue. Coverage is discussed in further detail in Section 3.5.1 and in Annex 3.

Third, a limitation should be noted with specific reference to criteria number six: "Using evaluation and monitoring to improve development effectiveness". Special care should be taken in interpreting the results in this area, as most of the evaluations reviewed do not directly address the strength of evaluation and monitoring as a part of their mandate to assess program effectiveness. The evaluations' authors do refer to these systems, but more in passing than in a direct effort to evaluate them. To some extent, the negative findings may reflect a bias in reporting on the part of evaluators because they are more likely to note weaknesses in the data they work with. Thus, it is possible that those evaluations that do not address monitoring and evaluation had positive findings in this area. This bias of non-response is much less likely in the other five criteria areas since they were normally a part of the mandate of the evaluation teams.

Fourth, like all evaluation reports, the evaluations reviewed are retrospective in nature. While published between 2009 and 2011, they describe programming realities that were in effect from as early as 2004-2005. All of the evaluations covered at least some programming between 2009 and 2011, and most evaluations focused on the recent program periods. In spite of this, some recent initiatives undertaken by UNDP to improve development effectiveness were not captured by the review, because of the retrospective nature of evaluations.

3.0 Findings on the Development Effectiveness of UNDP

This section presents the results of the review as they relate to the six major criteria and their associated sub-criteria (See Table 2 and Annex 1). In particular, Table 2 below describes the proportion of evaluations reporting findings of "satisfactory" or "highly satisfactory" for each of the six major criteria and their associated sub-criteria. The table also presents the number of evaluations that addressed each sub-criterion (represented by the letter *a*).

Each of the following six sections begins with a summary of key findings, including areas where UNDP has made a significant contribution. Subsequently, the report describes the importance of positive and negative factors contributing to results under each assessed criteria of development effectiveness by quantifying how many evaluations identified a particular factor.

The following sections cite examples drawn from the reviewed evaluations. The numbers referenced in these examples refer to the frequency with which analysts in the review team highlighted these policy and program initiatives when reviewing evidence to support the findings in a given evaluation. However, contributing factors were not always explicitly reported in all reviewed evaluations. Thus, figures cited do not provide an exhaustive census of the contributing factors. In spite of this, the references do present an overall portrait of the emphasis given in different evaluation reports to the various factors.

3.1 Evaluations report that most UNDP programs achieve their development objectives and expected results

3.1.1 Coverage

The four sub-criteria relating to objectives achievement were addressed in most evaluations (see Figure 2). Sub-criterion 1.1 (programs and projects achieve stated objectives) was addressed by all 55 evaluation reports in the sample. Similarly, sub-criteria 1.2 and 1.4 were rated "high" in coverage, with only 8 and 7 evaluations, respectively, not addressing these sub-criteria. Sub-criterion 1.3 was rated "moderate" in coverage, as 12 evaluations did not address the number of beneficiaries and contribution to national goals.

Figure 2: Number of Evaluations Addressing Sub-Criteria for Objectives Achievement (n=55)[21]

---
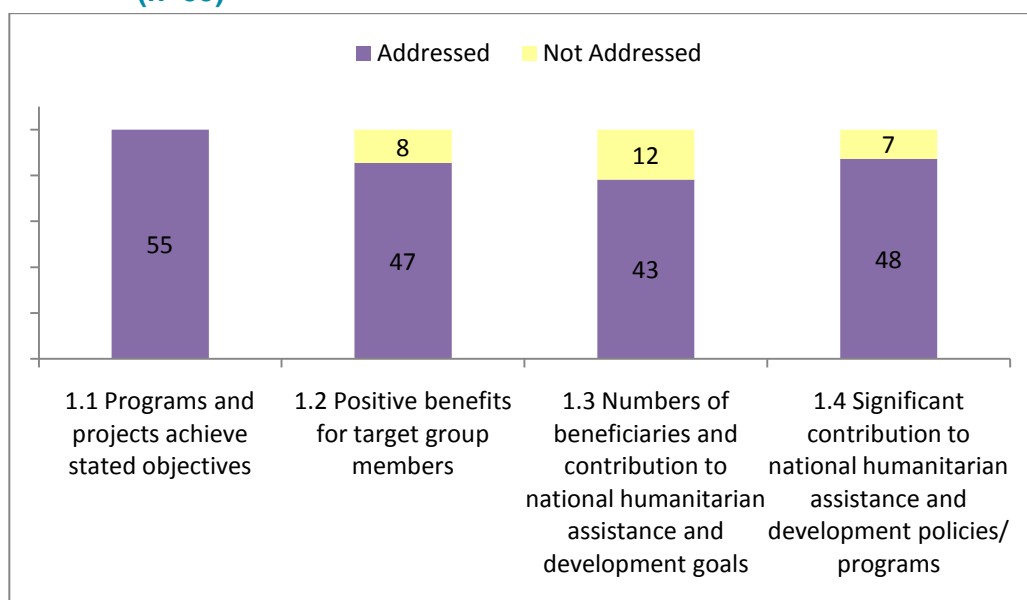
[21] *n* = the number of evaluations in the sample.

Table 2: Percentage of Evaluations Reporting Findings of "Satisfactory" or "Highly Satisfactory," per Criterion and Sub-criterion

| Criterion 1: Achieving Development Objectives and Expected Results	%*	a**
1.1 Programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.	66	55
1.2 Programs and projects have resulted in positive benefits for target group members.	83	47
1.3 Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.	65	43
1.4 Programs contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	71	48
Criterion 2: Cross Cutting Themes – Inclusive Development which is Sustainable	**%***	**a***
2.1 Extent to which multilateral organization supported activities effectively addresses the crosscutting issue of gender equality.	62	50
2.2 Extent to which changes are environmentally sustainable.	79	38
Criterion 3: Sustainability of Results/Benefits	**%***	**a***
3.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations to rehabilitation, reconstructions and, eventually, to longer term developmental results.	36	50
3.2 Projects and programs are reported as sustainable in terms of institutional and/or community capacity.	58	53
3.3 Programming contributes to strengthening the enabling environment for development.	75	49
Criterion 4: Relevance of Interventions	**%***	**a***
4.1 Programs and projects are suited to the needs and/or priorities of the target group.	73	53
4.2 Projects and programs align with national development goals.	88	51
4.3 Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	77	53
Criterion 5: Efficiency	**%***	**a***
5.1 Program activities are evaluated as cost/resource efficient.	52	39
5.2 Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).	51	31
5.3 Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).	27	41
Criterion 6: Using Evaluation and Monitoring to Improve DE	**%***	**a***
6.1 Systems and process for evaluation are effective.	41	42
6.2 Systems and processes for monitoring and reporting on program results are effective.	23	48
6.3 Results based management systems are effective.	24	37
6.4 Evaluation is used to improve development/humanitarian effectiveness.	81	47

* Percent of evaluations reporting findings of "satisfactory" or "highly satisfactory."
**a = Number of evaluations that address the given sub-criteria.

3.1.2 Key Findings

The reviewed evaluations report that most UNDP programs are achieving their development objectives and expected results (see Figure 3), although weaknesses remain in some areas. Regarding UNDP's performance on achieving its objectives, 66% of evaluations rated UNDP programs as "satisfactory" or better (1.1). Most (83%) evaluations reported findings that were "satisfactory" or better for positive benefits for target group members (1.2). Results were also positive for the number of beneficiaries reached by UNDP programs (1.3), with 65% of evaluations reporting "satisfactory" or "highly satisfactory" findings. Finally, influencing changes in national development policies and programs (1.4) was reported "satisfactory" or "highly satisfactory" by 71% of evaluations.

The review did not assess differences between the objectives achievement in the four pillars of UNDP's organizational mandate (i.e. poverty eradication, environment and sustainable development, democratic governance and crisis prevention and recovery). Nonetheless, the evidence cited in this section illustrates positive development results from UNDP programming.

Figure 3: Achievement of Development Objectives and Expected Results (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)[22]
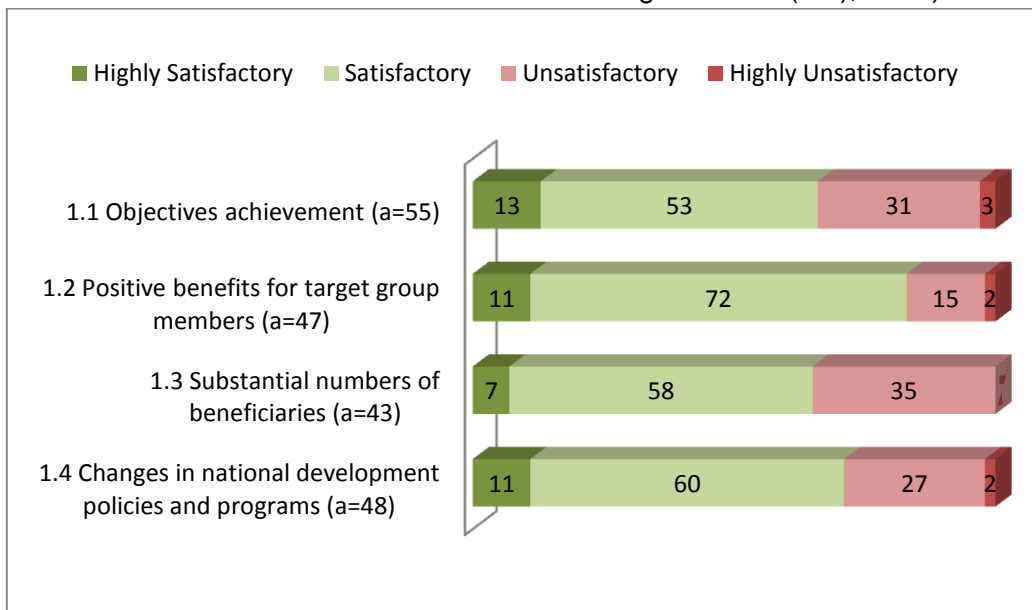
The reviewed evaluations emphasized the types of benefits accruing to target group members. The benefits included increased capacity resulting from UNDP supported training and knowledge development efforts. Highlights include:

-   Enhanced capacity due to training and capacity development activities that strengthened the transparency, responsiveness and probity of governance. Results include improved professional capacity for the police, the legal profession, judiciary, and parliamentarians (18 evaluations).

-   Improved livelihoods resulting from microfinance services, income generation and livelihoods development (15 evaluations).

---

[22]*n* = the number of evaluations in the sample. *a*= the number of evaluations that addressed the sub-criteria.

- Improved capacity of public servants in development planning, administration, management and evaluation (10 evaluations).

- Improved capacity of community leaders and civil society members in environmental management for poverty reduction (7 evaluations).

- Better health management by people living with HIV because of counseling and training combined with other therapies (5 evaluations).

- Improved capacity for disaster management and risk reduction (4 evaluations).

UNDP influence on national government development policies and programs mostly involved improved democratic governance (see Highlight Box 1). Areas where UNDP has made a significant contribution to changes in policy include:

- Governance reform, including increased transparency, strengthened parliamentary systems, improved judicial and policing systems, and enhanced peace-building (16 evaluations).

- Increased use of the MDGs to focus programs and policies on pro-poor development, including human development reporting, poverty mapping, poverty assessment and the development of statistical systems for targeting poverty (10 evaluations).

- Improved national policies and programs on disaster management and national disaster risk reduction strategies and programs (8 evaluations).

- Strengthened national decentralization policies and improved policies, systems and procedures at local levels (7 evaluations).

- Increased attention and awareness among policy makers of links between environment and development, including natural resource management, energy policy and climate change policy (7 evaluations).

- Strengthened national policies on gender mainstreaming and gender analysis, including human rights-based approaches to gender equality and improved programs to combat human trafficking and domestic violence (6 evaluations).

Highlight Box 1:
Using Knowledge Products to Achieve Development Objectives (Philippines)

"[The governance program] was prolific in the development of knowledge products, ranging from citizens' guides to monitoring government, to the *Philippine Governance Forum* (a series of public forums on key governance themes and issues), to a comprehensive manual on a *Rights Based Approach* (RBA) to development. The cluster approach, involving a wide range of participants, proved a good mechanism for disseminating these products."

Philippines ADR, 2009, p. 32.

3.1.3 Contributing Factors

The evaluations reviewed identified several factors contributing to or hindering UNDP's achievement of objectives. Positive factors cited in this area include:

- UNDP's investment in the creation and dissemination of knowledge products to support policy development by government. This support also included advocacy for the priorities of poverty reduction, gender equality and environmental sustainability (11 evaluations).

- UNDP's practice in some programs and countries of acting to bring together government and civil society at national and local levels, and through consultations and constituency building to improve program design and build support for priority areas (6 evaluations (see also Highlight Box 2).

- UNDP's ability to advocate for the MDGs and their integration into policies and programs. UNDP's role was noted as a means of increasing national and local government support for programming in these areas (6 evaluations).

- A factor in successful changes in national development policies was national policy support linked with direct local interventions, thereby linking lessons learned locally with national policies (4 evaluations).

- UNDP's organizational strengths in analysis, planning, and management and its strong commitment to capacity development and training was a positive factor for benefits reaching target group members.

- In addition, UNDP's reputation for reliability based on its unique position among donors and national actors was a positive factor cited for governance reform.

Factors cited which hindered the achievement of program objectives include:

- In some countries, UNDP programming was spread across too many projects, too large a geographic area, or too many institutions/ target groups. The dispersion was often made worse by overly ambitious project and program goals or limited financial and human resources (11 evaluations).

- UNDP did not always take full advantage of its strategic position and strong reputation (based on programming strength) to maximize its potential policy development impact (4 evaluations).

- UNDP's efforts to engage in improving democratic governance through enhanced transparency and reduced corruption were met by indifference or even hostility on the part of national governments (3 evaluations).

Highlight Box 2:
Contributing to Local Governance Reform in Indonesia

"The Aceh Justice Project supports seven civil society organizations that assist poor claim holders. It has also conducted legal awareness campaigns and developed training material for the formal and informal justice systems. It demonstrates that significant contributions can be made to governance reform with relatively small resource inputs, provided that they are sharply focused on problems of strategic importance."

Indonesia ADR, 2010, p.24.

3.2 UNDP works to mainstream gender equality and promotes environmental sustainability

3.2.1 Coverage

UNDP evaluations provided a high level of coverage of gender equality (2.1), with only five evaluations coded as "not addressed" (see Figure 4). For environmental sustainability (2.2), coverage was moderate, with 17 evaluations not addressing this issue.

Figure 4: Number of Evaluations Addressing Sub-Criteria for Inclusive Development (*n*=55)
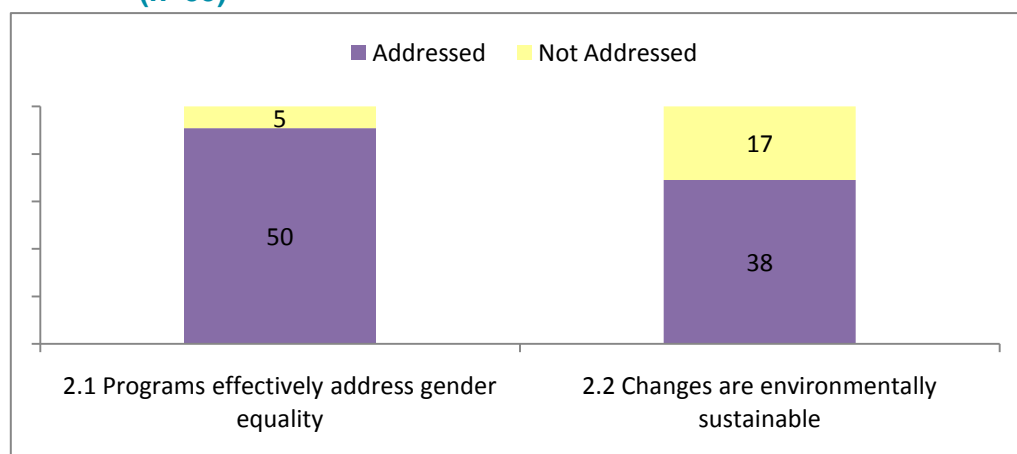
3.2.2 Key Findings

A majority of the evaluations (62%) rated programs as "satisfactory" or "highly satisfactory" in support for gender equality (2.1). Many UNDP programs have mainstreamed gender equality into program components. They have also helped to increase women's participation in governance structures. However, 38% of evaluations rated gender equality results as "unsatisfactory" or "highly unsatisfactory." The most frequently cited factor limiting effectiveness in gender equality was the absence of a systematic approach to gender analysis during program design.

UNDP's effectiveness in supporting environmentally sustainable development (2.2) was assessed as positive. This area ranks third highest of all the sub-criteria used to measure development effectiveness, in terms of percentage of evaluations reporting "satisfactory" or "highly satisfactory" findings (79%).

Figure 5: Inclusive Development which can be Sustained (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)
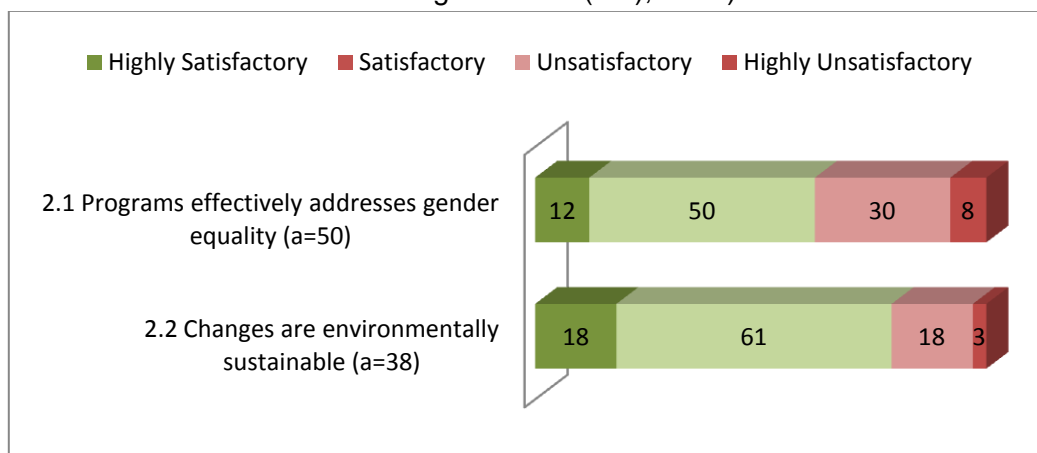
Evaluations that reported effective UNDP support to gender equality noted that UNDP had mainstreamed gender equality into its four main program areas: poverty reduction; environment and sustainable development; democratic governance; and, crisis prevention (16 evaluations). In particular, these evaluations noted success in:

– Strengthening women's roles in governance at a national and local level (10 evaluations).

– Development and distribution of knowledge products on gender equality including the use of sex-disaggregated data in reports, web sites, brochures, etc. (6 evaluations, see Highlight Box 4).

– Advocacy and policy support to the Convention on the Elimination of All forms of Discrimination against Women (CEDAW) and for the gender equality targets of the MDGs (6 evaluations).

The most frequently cited successful areas of UNDP programming for environmental sustainability include:

– Enhanced national energy policies, the introduction of measures to improve energy efficiency, and efforts to develop a national climate change strategy (8 evaluations).

– Improved rural and urban water resource management (4 evaluations).

– Strengthened conservation programs and improved promotion of bio-diversity (4 evaluations).

– Improved natural resource management capacity, including capacity for sustainable land management (4 evaluations).

– Support to fulfill national obligations arising from multilateral agreements (2 evaluations, see Highlight Box 3).

Development Effectiveness Review of the United Nations Development Programme

Highlight Box 3:
Supporting Environmental Sustainability

"UNDP regional programs and projects have consistently supported program countries in addressing common environmental issues. An example of UNDP contribution to the achievement of shared results includes the drafting of the Nile River Basin cooperative framework in the Arab region, which was supported through the financing of technical studies and the facilitation of the subsequent intergovernmental dialogue. Another example is the operationalization of an information network that monitors and shares regional practices in toxic waste management in the Mekong River Basin countries, which was developed with UNDP support."

Evaluation of UNDP Contribution at the Regional Level to Development and Corporate Results, 2010, p. 29

Highlight Box 4:
Contributing to Gender Equality in the West Bank and Gaza

"Women have benefited from many of Program of Assistance to the Palestinian People (PAPP) projects, including the Introduction to Participatory Planning Programme, the KFW and Land Reclamation Programme and the first phase of the Leadership Development Programme. After the Gaza incursion, specific attention was given to the needs of women through an immediate focus group discussion with three women leaders representing a women's umbrella organization. This provided sufficient information for the design of a Social Development Assessment that is examining a broad spectrum of questions about social inclusion, targeting men and women, old, young, disabled and healthy, marginalized and geographically central Gazans.

"UNDP, in cooperation with the Palestinian Agricultural Relief Committees (PARC), has supported the Rural Women's Credit and Savings Association. The project began in 1999, as a saving and credit programme to provide economic assistance to members of a women's cooperative in the West Bank. Over the years, PARC expanded its work to include capacity building based on international best practices. The women's cooperative, now oversees the management of funds, with PARC continuing to provide technical and logistical support. The project is ongoing in 132 locations in the West Bank and Gaza with 5,653 rural women.

"UNDP was also part of the interagency team that secured a grant from the Spanish MDG Achievement fund in the window on 'women', and is expanding its work on the prevention and response to sexual and gender-based violence."

Outcome Evaluation of the UNDP / Program of Assistance to the Palestinian People Mid-Term Strategic Framework 2008-2011, 2009, p.xii.

3.2.3 Contributing Factors

As for all sub-criteria, a number of the factors contributing to or detracting from the effectiveness of programs in addressing sustainability are not fully within UNDP's control. In

particular, addressing some of the negative factors cited here would require improved performance by UNDP's partners.

Positive contributing factors for promoting inclusive development in the areas of gender equality and environmental sustainability include:

- The combination of successful women-specific programs at local and national levels, with a policy of advocating for gender mainstreaming which is incorporated into UNDP supported programs (11 evaluations).

- Development and use of research and knowledge products on gender equality to support advocacy and policy dialogue with government and civil society (4 evaluations).

- Strong and consistent advocacy for effective gender equality programming (3 evaluations).

- Effective advocacy by UNDP to keep environmental issues and concerns at the top of the national policy agenda (4 evaluations).

- UNDP's ability and willingness to invest in environmental research methods and tools, and to develop and disseminate knowledge products (4 evaluations).

Factors that limited effectiveness in gender equality and environmental sustainability include:

- Poor conceptualization at the local level of a gender equality model, which can be implemented in the local context without weakening the commitment or results achieved (11 evaluations).

- The lack of human resource capacity for gender programming in some UNDP offices (4 evaluations).

- Poor conceptualization of gender equality so that it can be applied to the local context (3 evaluations).

- An overly diverse set of environmental projects in a single country (especially for energy programs), lacking in an overall strategic focus (3 evaluations).

- Insufficient involvement by civil society in environmental sustainability issues (2 evaluations).

- Programs not always addressing some of the most important national challenges to environmental sustainability, for example land degradation and desertification. (2 evaluations).

3.3 The sustainability of benefits is a significant challenge

3.3.1 Coverage

All three sub-criteria for sustainability have high levels of coverage, with most evaluations addressing these issues. As illustrated by Figure 6, very few evaluations were coded as "not addressed" for each of the three sub-criteria in this area.

Figure 6: Number of Evaluations Addressing Sub-Criteria for Sustainability (*n*=55)
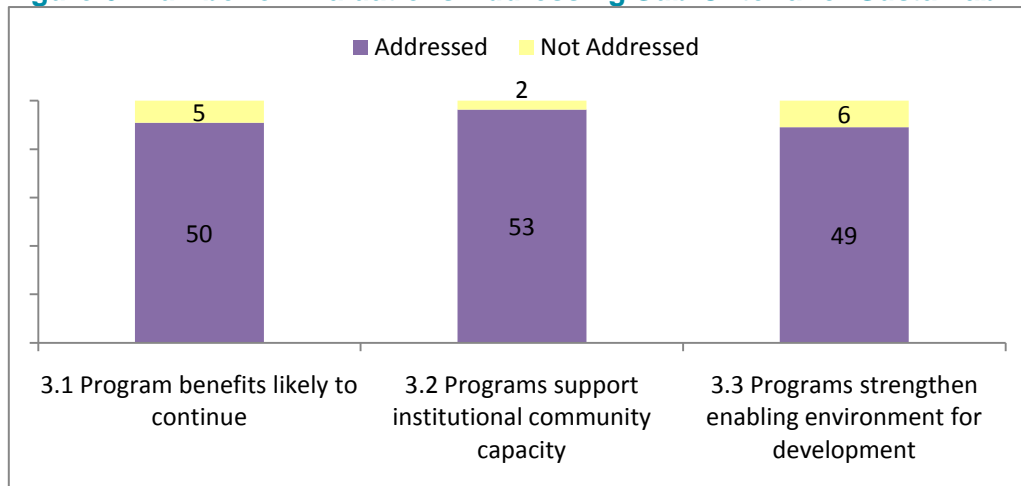
3.3.2 Key Findings

The sustainability of results and benefits represents a significant challenge to the development effectiveness of UNDP. Only 36% of the evaluations reviewed assessed the likelihood of benefits continuing in the post-program period as "satisfactory" or better (3.1). This was the fourth lowest ranking sub-criteria measuring development effectiveness.

The evaluation findings on institutional and community aspects of sustainability (3.2) are more positive, with 58% of evaluations reporting "satisfactory" or "highly satisfactory" results. However, significant weaknesses remain in this area, as 21% of the reviewed evaluations reported that UNDP programs were "highly unsatisfactory" in their support to institutional and community capacity for sustainability.

UNDP achieved positive results on the extent to which UNDP programs have made a positive contribution to the enabling environment for development, mainly through their influence on national development planning, budgeting and programming (3.3). Most of the reviewed evaluations (75%) reported results that are "satisfactory" or "highly satisfactory" in this area.

Figure 7: Sustainability of Results/Benefits (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)
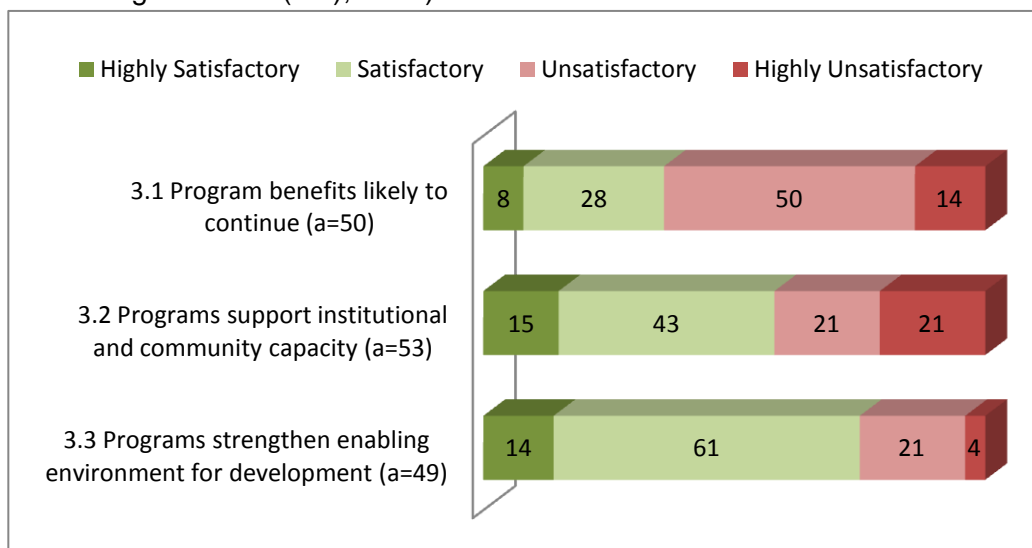
Evaluation reports highlighted the types of support that contributed most directly to strengthening the enabling environment for development at national and local levels. These types of support include:

- Direct support to improved democratic governance (18 evaluations).

- Strengthening civil society organizations and facilitating participation by civil society in national and local policy and programming (17 evaluations).

- Improved data collection methods and techniques for mapping poverty and living standards, and for supporting research used to develop and promote pro-poor development planning (8 evaluations).

- Direct support to decentralization (7 evaluations).

- Supporting improvements in government capacity to coordinate and manage development assistance within the framework of the Paris Declaration (6 evaluations).

- Advocacy for MDGs and pro-poor development policies (5 evaluations).

3.3.3 Contributing Factors

Positive factors contributing to the sustainability of benefits include:

- Strong program ownership by the developing country partners at all levels: national governments, local government, communities and civil society organizations (7 evaluations).

- The use of participatory planning processes to develop local ownership and identify sustainable co-financing strategies as a positive factor in the sustainability of anti-poverty programs (4 evaluations).

- The integration of UNDP supported programs into the national budget process and the Medium Term Expenditure Framework (3 evaluations).

Development Effectiveness Review of the United Nations Development Programme

Factors contributing to weakened sustainability of benefits include:

- The absence of an explicit phase-out strategy and integration of sustainability in the early stages of program design and implementation (14 evaluations, see Highlight Box 5).

- Inadequate or interrupted funding from external sources, particularly near the end of programs (10 evaluations).

- A high level of dependence on UNDP during program implementation, as evidenced by weak government commitment and a lack of ownership at national and/or local levels (9 evaluations).

- Lack of needs-assessment at the institutional and community level. As a result, capacity development programs were inadequate and overambitious (5 evaluations).

- Non-integration of the costs of program activities into the national budget and the Medium Term Expenditure Framework (i.e. inadequate provision for recurrent costs by partner countries) (4 evaluations).

- Program design included sustainability elements, but these elements were not delivered on a continuous basis or were delivered too close to the end of the program, when resources were declining and national government interest was limited (3 evaluations).

Highlight Box 5:
Challenges to Sustainability in Ghana

"Sustainability is a challenge across UNDP Ghana's programme portfolio, though it plays itself out slightly differently in the various thematic areas and across different types of projects. Although all Annual Work Programs contain details of threats to sustainability, the risk mitigation strategies are seldom in place or followed consistently. Explicit exit strategies are the exception rather than the rule."

Ghana ADR, 2011, p.49.

3.4 UNDP programs are relevant to the context in which they work

3.4.1 Coverage

Sub-criteria relating to relevance were addressed by most of the reviewed evaluations (see Figure 8). These three areas all have "high" levels of coverage.

Figure 8: Number of Evaluations Addressing Sub-Criteria for Relevance (*n*=55)
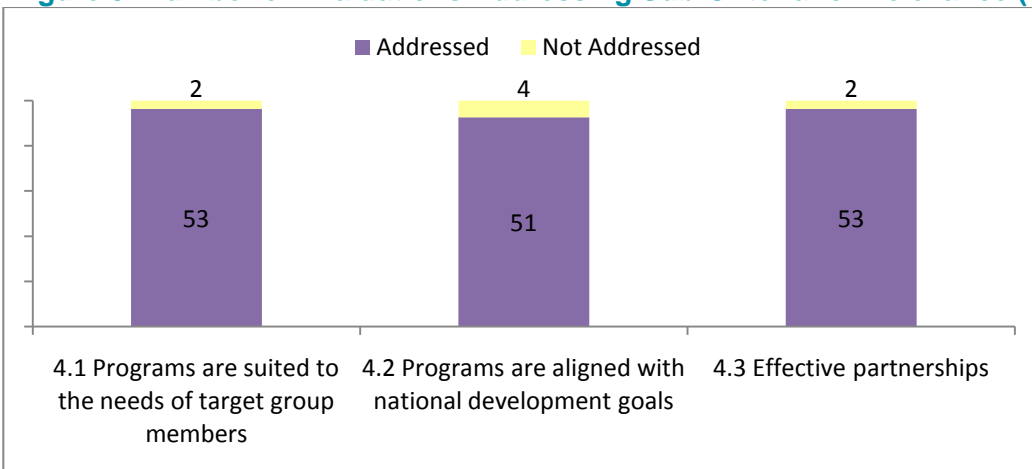
3.4.2 Key Findings

Relevance was assessed positively in the reviewed evaluations. The suitability of UNDP programs to the needs of target group members (4.1) was rated as "satisfactory" or better in 73% of the evaluations reviewed. Evaluation findings were also positive regarding the alignment between national development goals and UNDP supported programs (4.2), with 88% of evaluations reporting "satisfactory" or "highly satisfactory" results. 77% of evaluations reported positive findings for sub-criterion 4.3 on effective partnerships with government, civil society and development partners (see Figure 9).

Figure 9: Relevance of Interventions (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)
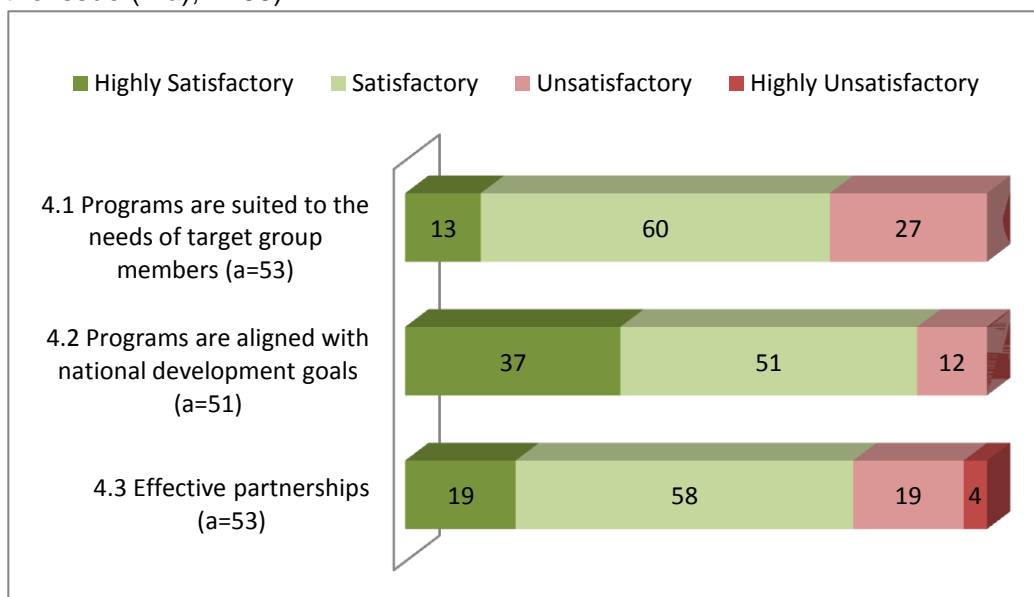
Evaluation reports highlighted some types of programs that were especially relevant to the needs of target group members. They pointed to programs for sustainable livelihood development, poverty reduction, the nexus between poverty and the environment, improved gender equality, and improved access to justice as good examples of a close fit between individual needs and program outputs (16 evaluations). Evaluation reports also emphasized the positive alignment between UNDP supported programs and national development priorities and needs, usually expressed in terms of the national development plan (16 evaluations, see example in Highlight Box 6).

The evaluation reports also identified three types of successful partnerships established by UNDP:

– UNDP has formed strong partnerships with civil society organizations and has worked in many countries to facilitate their participation in national and local development planning and programming. UNDP has worked to develop networks, meetings and forums that would create a space for participation by civil society organizations in the dialogue on development policies and programs (22 evaluations).

– UNDP has formed effective partnerships with key agencies of national and local governments. These partnerships involve collaboration on policies and priorities, effective advocacy relationships, support to research on pro-poor policies and programs, and support for national development planning processes (21 evaluations).

– UNDP also participated in forums on donor coordination, national development policies and programs, including the United Nations Development Assistance Framework,[23] the Joint Assistance Strategy,[24] joint government-donor sector working groups, and annual public consultations. These forums encompass national governments, bilateral donor partners, multilateral development banks and other organizations of the UN (20 evaluations).

Highlight Box 6:
Malawi Programs Rated Highly Relevant

"The ADR has demonstrated that the country programme is highly relevant to national development priorities and the overall mandate of promoting sustainable human-centred development underpinned by poverty reduction, equity, fairness and justice. In addition, UNDP has demonstrated a great deal of responsiveness to the country's changing context, by adjusting its programme portfolio to reflect national development priorities. UNDP has also demonstrated its responsiveness to emerging needs, including the need to strengthen human capacities. Its programmes have paid particular attention to the human development dimensions of gender equality, women's rights and vulnerable groups."

Malawi ADR, 2011, p. 53.

---

[23] The UNDAF details the ways in which the UN family will support the national development plan. It includes objectives for each UN agency that must contribute to overall UNDAF results and has a results framework at the UN system level.
[24] The Joint Assistance Strategy (JAS) is a political document signed at the ambassadorial level governing how the bilateral donors and the multilateral agencies will contribute to national development goals in some countries.

3.4.3 Contributing Factors

Positive factors contributing to the relevance of interventions include:

–   A systematic approach to consultations among all stakeholders (state and civil society) in order to reach consensus and a high level of agreement on specific needs of target group members and the best means to address them (11 evaluations).

–   UNDP's strategy of advocating for the MDGs and pro-poor policy perspectives in national development planning processes. This allows linking national and global development priorities and permits UNDP to enter into programming areas where it has the most experience (8 evaluations).

–   The ability of some host governments to develop a clear, well articulated set of national development priorities that are backed up by policy documents and guidelines (7 evaluations).

–   Conducting problem analysis at the national and local levels, especially in the design of livelihoods programming (5 evaluations).

–   Success in developing effective partnerships was attributed to the ability of the UNDP country office to develop a reputation for neutral and unbiased policy advice and program support. This allows UNDP to develop coalitions of partners with different policy perspectives and work with them to develop consensus (5 evaluations).

–   The ability of some UNDP offices to play an important role in facilitating the development of national priorities and programs (4 evaluations).

–   UNDP's flexibility in adapting to emerging national needs, including humanitarian needs and changes in conflict situations (3 evaluations).

Factors contributing to negative ratings in the relevance of interventions include:

–   Lack of a systematic approach to conducting needs assessments during the planning of some UNDP-supported programs, including the absence of a systematic approach to risk assessment (11 evaluations). This contributed to a poor fit between programs and the needs of target group members.

–   Effective partnerships were limited by a lack of active engagement with civil society organizations in some countries, failing to support their integration into the development of national and local policies and programs, which, in turn, weakens the consensus on priority needs and solutions (12 evaluations).

–   Effective partnerships were also limited by poor coordination with bilateral donors and other UN agencies, sometimes because of an environment of inter-agency competition (7 evaluations).

3.5 Evaluations report weaknesses in program efficiency

3.5.1 Coverage

All three sub-criteria for efficiency were rated "moderate" in coverage (see Figure 10). 16 evaluations did not address sub-criterion 5.1 (programs are cost efficient) while 24 did not address 5.2 (programs implemented on time) and 14 evaluations failed to address sub-criterion 5.3 (systems for program implementation are efficient).

Figure 10: Number of Evaluations Addressing Sub-Criteria for Efficiency (n=55)
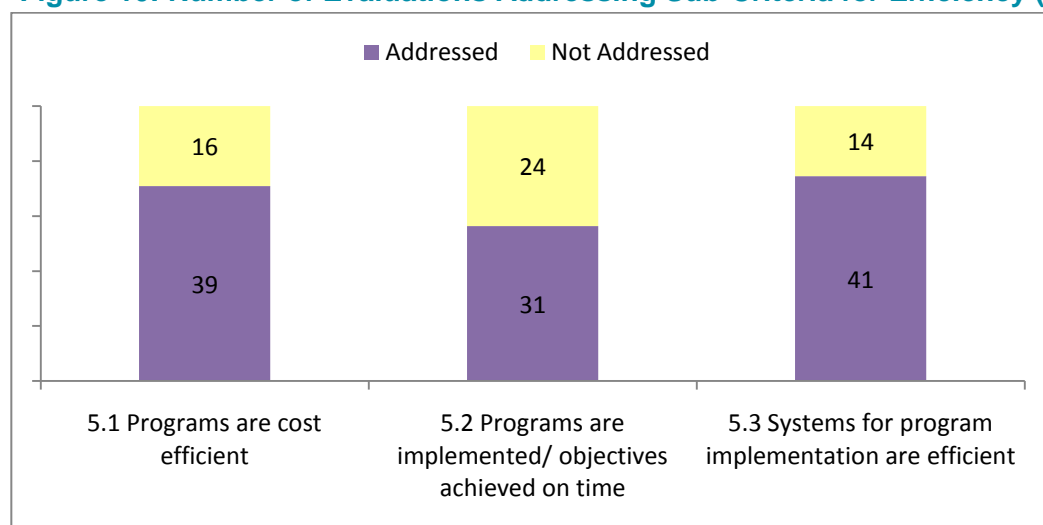
Based on the coverage of these sub-criteria, it is important to treat the results reported for efficiency with some care. Sub-criteria 5.1 and 5.2 earned a "moderate" coverage assessment, based on the numbers of evaluations that did not address the given issues. Nevertheless, this is still a solid base of evidence from which to draw conclusions regarding the development effectiveness of UNDP.

3.5.2 Key Findings

The cost efficiency of UNDP programs was rated "satisfactory" or better in only 52% of the evaluations reviewed. Results were similar regarding the timeliness of program implementation, with 51% of evaluations reporting "satisfactory" or "highly satisfactory" findings.

During the review process, an item was added regarding the efficiency of UNDP systems for administration and control of programs in the field (5.3), because the evaluations reviewed included many findings related to administrative systems as a factor in UNDP's development effectiveness. This item was found "satisfactory" or better in only 27% of 41 evaluations that addressed this issue. This was the most negative finding for any of the 19 sub-criteria. Findings on efficiency in the reviewed evaluations indicate that this is one of the most challenging areas requiring improvements in order to further strengthen UNDP's development effectiveness.

Figure 11: Efficiency (Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)

Legend: ■ Highly Satisfactory ■ Satisfactory ■ Unsatisfactory ■ Highly Unsatisfactory

5.1 Program activities are cost efficient (a=39): 3 | 49 | 38 | 10

5.2 Programs are implemented/objectives achieved on time (a=31): 3 | 48 | 39 | 10

5.3 Systems for program implementation are efficient (a=41): 27 | 61 | 12
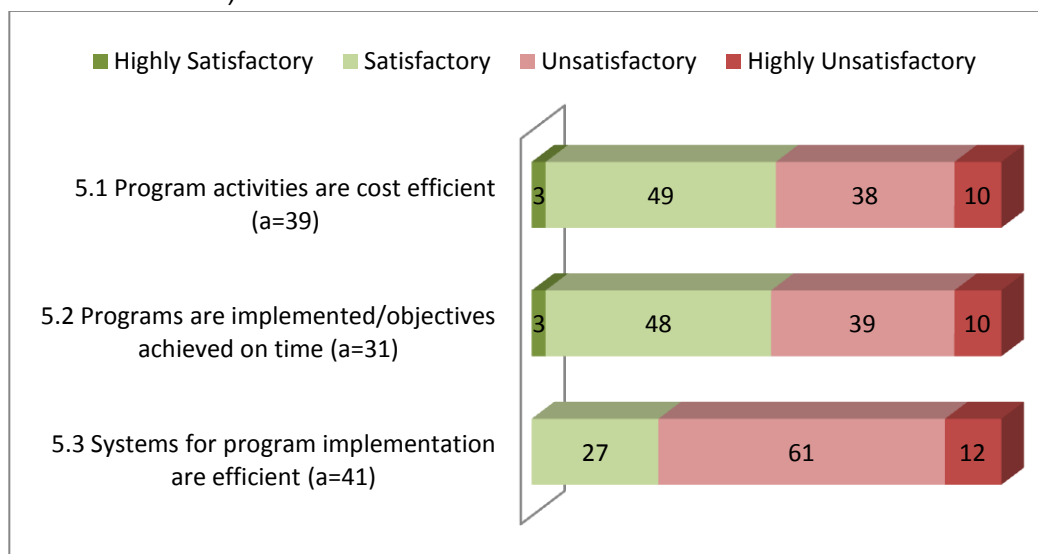
3.5.3 Contributing Factors

Only one factor was identified in more than one evaluation as contributing positively to program cost efficiency. It concerns UNDP's ability to deliver programs within a constrained budget and with limited financial resources of its own. UNDP achieved this objective in some programs by leveraging outside resources and implementing low-cost program elements (5 evaluations).

The challenges in achieving efficiency in UNDP programs and systems were highlighted by an array of factors contributing to negative findings in this area, including:

– The complexity and rigidity of UNDP systems, processes and procedures for project administration and control, particularly in relation to procurement and the disbursement of funds (16 evaluations, see Highlight Boxes 7 and 8).

– Program resources spread too thinly across too many projects in diverse geographic locations. Consequently, administrative costs rise, as costly administrative functions are duplicated across too many small projects (8 evaluations).

– Staff shortages and excessive turnover in UNDP country offices (8 evaluations, see Highlight Box 8)

– The cost of compliance with UNDP global regulations and administrative frameworks governing procurement, finances, service contracts, security, and travel (5 evaluations).

– The challenge to deliver results on time because of project and programs being designed with overly ambitious objectives (5 evaluations).

– Lack of partner capacity for program delivery in accord with the program timetable. This problem was made more difficult by a tendency to underestimate the capacity development challenge in developing program designs (5 evaluations).

– Excessive and overly complex reporting requirements (4 evaluations).

– Delayed, erratic or uncertain funding. This results in significant increased costs as programs are stopped and started to match funding (3 evaluations).

Highlight Box 7:
Inefficient Administrative Procedures in Afghanistan

"Flexibility and rapidity of response are crucial in conflict-affected countries. Yet if there is one universal criticism of UNDP from stakeholders in Afghanistan, it is their devastating indictment of the inefficiency of UNDP bureaucratic procedures. So much ill will has been created as a result of massive delays in procurement, payments and other basic administrative tasks it threatens to overwhelm the substantive achievements of the program.

"Steps have been taken during the past two years to improve administration, but the fundamental problem of inefficiency and procedural complexity is systemic. There are no special financial, procurement and human resources guidelines for the needs of large post-conflict country offices. The introduction of new financial asset software has further added to the inflexibility, and the system appears to be getting more rigid and bureaucratic rather than less."

Afghanistan ADR, 2009, p. xvii

Highlight Box 8:
Slow Procedures and High Staff Turnover Limit Efficiency in Indonesia

"While UNDP is usually very efficient in its dealings at the central level with other partners and Government of Indonesia agencies, its delivery of support for projects is regarded as slow and not cost effective. Long chains of authority delay decision making in the country office and, in particular, when Headquarters in New York is involved. A high turnover rate among project management staff has also been noted as a contributing factor."

Indonesia ADR, 2010, p. 25.

3.6 UNDP faces challenges in strengthening decentralized systems for evaluation, monitoring and results-based management

3.6.1 Coverage

Coverage was high for two of the four sub-criteria relating to the use of evaluation and monitoring to improve development effectiveness (see Figure 12). Sub-criterion 6.2 was addressed by 48 evaluations, while sub-criterion 6.4 was addressed by 47 evaluations (see Figure 12). The remaining two sub-criteria (6.1 and 6.3) were addressed by 42 and 37 evaluations, respectively.

Figure 12: Number of Evaluations Addressing Sub-Criteria in Evaluation and Monitoring (*n*=55)

Legend: Addressed | Not Addressed

Sub-Criteria	Addressed	Not Addressed
6.1 Systems for evaluation are effective	42	13
6.2 Systems for monitoring results are effective	48	7
6.3 Results-based management systems are effective	37	18
6.4 Evaluation results are used to improve development effectiveness	47	8
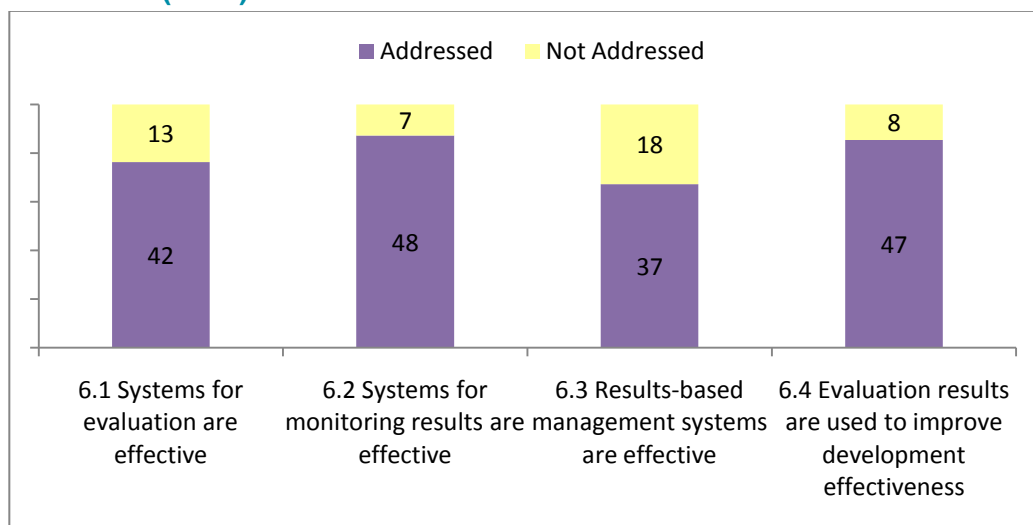
3.6.2 Key Findings

Findings reported under sub-criteria 6.1, 6.2 and 6.3 refer to the presence and strength of evaluation, program monitoring and results-based management at the country office level, as reported in the reviewed evaluations. They do not apply to corporate level evaluations managed by the UNDP Evaluation Office. The quality assessment of the 55 evaluations used in the review sample (as reported in Annex 4) resulted in positive ratings for all 55 evaluations. The reports in the sample were also notable for their frank presentation of the negative, as well as the positive, elements of UNDP's development effectiveness.

Nonetheless, UNDP faces challenges in strengthening systems for evaluation, results monitoring and results-based management at the decentralized program level. Evaluation systems and processes (6.1) were found to be effective at the "satisfactory" or "highly satisfactory" level in only 41% of the evaluations reviewed. The challenge of strengthening decentralized evaluations has been recognized by UNDP's Evaluation Office. In 2010, it reviewed a sample of decentralized outcome evaluations and reported that 36% were either "unsatisfactory" or "moderately unsatisfactory."[25]

Findings were even more negative for monitoring systems (6.2), with only 23% of evaluations reporting results of "satisfactory" or better. They were also negative for the effectiveness of results-based management systems (6.3), where only 24% of evaluations reported "satisfactory" or "highly satisfactory" findings.

Despite these negative findings for evaluation, monitoring and results-based management at decentralized levels, UNDP does make systematic use of evaluation findings to improve development effectiveness (6.4), with 81% of reviewed evaluations reporting that results were "satisfactory" or better.[26] In fact, 43% of the evaluations reported results that were "highly

---

[25] Annual Report on Evaluation, UNDP, 2011.

[26] For sub-criteria 6.4, the main indicator used by the review team was the frequency of a clear management response to the evaluations under review, as well as the adequacy, specificity and clarity of that response, especially where it includes an action plan with time bound commitments for taking action on recommendations. A "highly satisfactory" finding was coded when the

satisfactory" regarding the use of evaluation results to improve program effectiveness. This does not mean that poorly conducted or weak evaluations are being used to inform development programming at UNDP. Rather, the Evaluation Office, which is able to commission, carry out and follow up on quality evaluations, is providing a reasonably strong base of evaluation reports used to inform decision making.

Figure 13: Using Evaluation and Monitoring to Improve Development Effectiveness
(Findings as a % of number of evaluations addressing the issue (= *a*), *n*=55)
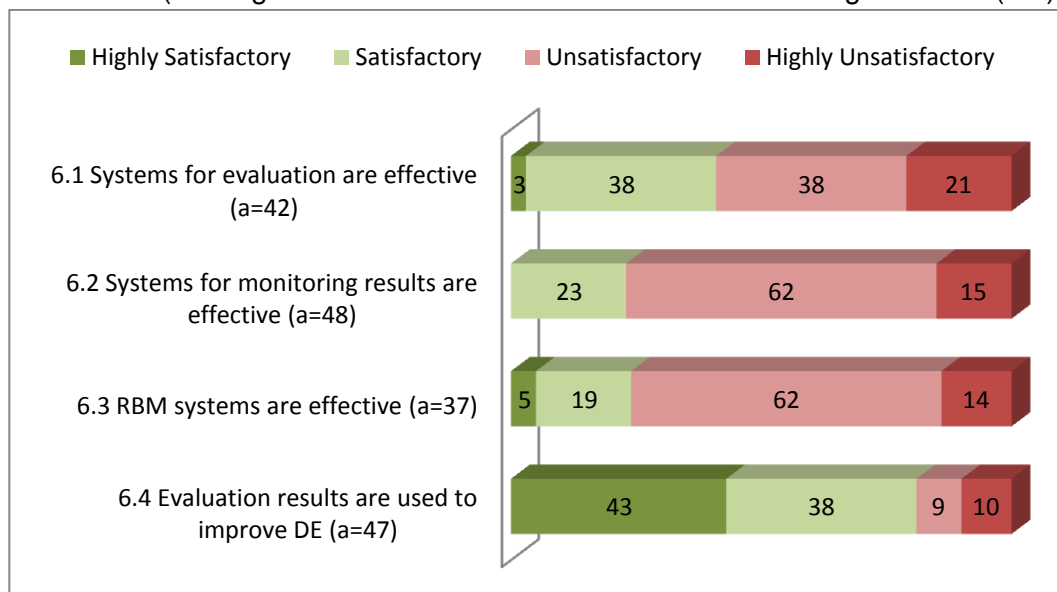
Highlight Box 9:
Weaknesses in Monitoring and Evaluation in the Ghana Program

"Robust monitoring and evaluation is a prerequisite for codifying lessons and generating knowledge from pilots and downstream work to influence upstream policy work. The weakness in monitoring and evaluation in the country office is not merely a reflection of the office's capacity constraints in this area. It is also a reflection of the serious limitations in monitoring and evaluation capacity of national partners. The quality of reporting from national partners is in many instances inadequate and the country office lacks the capacity to provide the necessary support and guidance."

Ghana ADR, 2011, p.59-60.

---

management response contained clear actionable items and some or all had been completed at the time of the Development Effectiveness Review.

Results on Evaluation and Monitoring

The negative findings on the strength of decentralized systems for evaluation do not mean that the evaluations used in the development effectiveness review of UNDP are of poor quality. First, 37 out of 55 evaluations in the sample were either Assessment of Development Results (ADRs) or global thematic evaluations carried out by UNDP's Evaluation Office (i.e. centralized evaluations).

Secondly, a meta-evaluation[1] was undertaken in order to ensure the quality of the evaluations used in this review. All 55 evaluations were assessed using the quality-screening tool (Annex 4). Each of the 18 decentralized evaluations received a score of 26 or higher from a possible 45. The mean quality score for the 55 evaluations in the sample was 37.7 of a possible 45.

Thirdly, in 2010, the Evaluation Office of UNDP reviewed a sample of 33 outcome evaluations and found 64% to be "satisfactory" or "moderately satisfactory."

In summary, the sample for this review consists of high quality evaluation reports produced by the Evaluation Office of UNDP and decentralized evaluations that have been screened for quality, providing a strong body of evidence for the review.

3.6.3 Contributing Factors

The evaluation reports reviewed tended to treat monitoring and evaluation as a single system and, as a result, cited the same contributing factors for these two areas. The most frequently cited hindering factors for monitoring and evaluation systems are:

- Weakness in results-based management was linked to the tendency to focus results models, indicators and reporting systems on program activities and outputs rather than the outcomes which really define program development effectiveness (11 evaluations).

- The absence of either evaluation studies or evidence-based monitoring reports for some specific national and regional projects or programs which were expected to inform decision-making and support centralized evaluations. For an example, see Highlight Box 10. (10 evaluations).

- Weak or missing baseline information to allow progress in securing results to be either monitored or evaluated (10 evaluations).

- The lack of adequate, trained human resources for monitoring and evaluation in UNDP country offices, sometimes combined with inadequate software and computerized data tools to support them (9 evaluations).

- Results-based management weaknesses were also attributed to the generally poor quality of the results framework developed during program design, with poorly specified indicators of outputs and outcomes and little understanding of the linkages between the two (7 evaluations).

- Partner reports that are deficient in reporting on program outcomes (5 evaluations).

- Results-based management systems were also weak due to a lack of a reporting and managing for results culture in some UNDP country offices and, especially, in the offices of implementing partners (3 evaluations).

- Lack of adequate funding for monitoring and evaluation (3 evaluations, see Highlight Boxes 10 and 11).

- Weak and poorly defined outcome indicators (3 evaluations).

- Lack of a comprehensive monitoring and evaluation strategy for regional or country programs (3 evaluations).

Highlight Box 10:
Lack of Project Evaluations for use in a Regional Assessment

"In terms of documentary evidence, the evaluation team had few projects evaluations and no outcome evaluations to work with. These decentralized evaluations are the building blocks of a regional program evaluation. Without them, collecting evidence across a large number of projects is extremely difficult."

Evaluation Mid-term Assessment of the Regional Program in Europe and the Commonwealth of Independent States, 2010, p. 4.

Highlight Box 11:
Limited Funds for Monitoring and Evaluation in Turkey

"While the office has developed best practices for monitoring and evaluation (e.g., for the regional development initiatives), corporate UNDP monitoring and evaluation practices generally are less comprehensive than those adopted by other multilateral organizations such as the World Bank or the European Commission. This is largely explained by the lack of adequate funding for project monitoring and evaluation. The country office, for example, does not have a monitoring and evaluation specialist. Weak monitoring and evaluation fails to do justice to UNDP's generally effective contribution to development results in Turkey. Thus, the many success stories cannot be substantiated with evidence from monitoring and evaluation, which threatens to constrain UNDP's partnerships with organizations with stronger monitoring and evaluation requirements."

Turkey ADR, 2010, p. 44.

4.0 Conclusions

The 55 evaluations reviewed provide reasonable coverage of a cross-section of recent UNDP's projects and programs. However, during the period under review (2005-2011), UNDP has continued to make changes to programming systems and procedures, including those under the 'Agenda for Change' initiative. Many of these changes respond to the findings of the evaluations used in this review. A comprehensive survey of the operational and organizational changes initiated by UNDP in the recent past is outside the scope of this review. However, UNDP has provided an overview of some of the most important of these recent changes (see Annex 7).

Based on the identified findings and related contributing factors, this review concludes that:

1. Evaluations used in this review indicate that most UNDP programs are achieving their development objectives and expected results, although weaknesses remain in some areas. The most frequently cited factor contributing to limited objectives achievement is dispersion of UNDP programming in one country across too many projects, too wide a geographic area or too many institutions. Positive factors contributing to UNDP's achievement of development objectives include effective investment in knowledge development; consultation to strengthen support for priority policies; and effective advocacy based on the MDGs.

2. UNDP is generally effective in addressing the cross cutting themes of gender equality and environmental sustainability. However, in order to improve the effectiveness of UNDP programs in gender equality, a formal, systematic approach to gender analysis should be applied during the design and implementation of all programs. In some countries, UNDP successfully supports gender equality through a strategy of combining women-specific programs at a local or national level, mainstreaming gender equality into all UNDP programs.

3. The sustainability of program results and benefits, represents a significant challenge to UNDP's development effectiveness. Reported factors limiting sustainability indicate that UNDP will need to increase efforts to ensure strong program ownership by developing country partners. It will also need to urgently address the absence in some programs of an explicit program phase-out strategy and the integration of sustainability into the early stages of program design. Improvements in sustainability will also require greater attention to assessing the capacities of partner country institutions to sustain program activities.

4. UNDP implements effective programs to strengthen the enabling environment for development. This involves directly supporting democratic governance; strengthening civil society participation in development planning; and, supporting government capacity for national development planning and programming.

5. UNDP programs are relevant to the needs of target group members and are aligned with national development priorities. The use of consultation among stakeholders to build a

consensus on specific needs and solutions was one factor contributing to programs that met the needs of target group members.

6. UNDP develops effective partnerships with civil society organizations to facilitate their participation in national and local development planning and programming. It also collaborates effectively with national governments on development policies and programs.

7. UNDP plays an important role in different forums for consultation on donor coordination and on policies and programs for development at a national level, including such mechanisms as the United Nations Development Assistance Framework, the Joint Assistance Strategy, joint government-donor sector working groups, and annual public consultations on development priorities at national level.

8. Improving the efficiency of UNDP programs at the country level represents a challenge for UNDP with many evaluations reporting unsatisfactory results. In order to improve the efficiency of UNDP programming, it will be necessary to address the problem of reportedly complex, rigid and overly bureaucratic UNDP systems, processes and procedures for project administration and control, particularly in relation to procurement and the disbursement of funds. In the area of efficiency, a necessary trade-off exists between flexibility and speed, on one hand, and accountability and transparency, on the other, when designing and implementing systems for administrative and financial project and program control.

9. The evaluations reviewed indicate that decentralized systems for monitoring and evaluation and results-based management are seriously deficient. However, the review also found that the evaluations produced by UNDP's Evaluation Office are of high quality and provide good coverage of UNPD programs and activity areas.

10. UNDP uses evaluation results to improve development effectiveness through a systematic and transparent process of detailed management responses containing actionable items. This use of evaluations to support development effectiveness relies on the work of the Evaluation Office, which conducts a significant number of high quality evaluations each year and works with decentralized offices to strengthen the quality of their evaluation outputs.

5.0 Areas for Attention

The previous section summarized the strengths and weaknesses in UNDP's development effectiveness. This section highlights the areas in which UNDP, working with its bilateral partners, needs to continue to focus attention.

1. Results on the achievement of objectives, while positive, also indicate that those programs that scored less than satisfactory for this criterion sometimes did so because of the dispersion of UNDP resources across too many projects, too wide a geographic area or too many institutions. There is a need for some UNDP programs to achieve greater focus by concentrating on fewer projects within a given country.

2. UNDP's program effectiveness in promoting gender equality is sometimes limited by inconsistent use of gender analysis in program design. UNDP needs to improve its effectiveness in mainstreaming gender equality into its development programming. This will require improved systems and processes for gender analysis during program design. It will also require strengthening the gender analysis capacity of country offices.

3. Given the review's conclusions on sustainability, UNDP should place a priority on implementing a systematic approach to developing explicit project phase-out strategies and sustainability designs that are integrated into the early stages of program development. In particular, there is a need to strengthen UNDP's quality at entry analysis, in areas such as risk analysis and needs assessment.

4. The review's conclusions on the efficiency of UNDP programming indicate the need to improve program efficiency at the country level. This would include priority attention towards improving the cost efficiency and timeliness of implementation of UNDP programs. In particular, UNDP systems and procedures for program and project administration and control need to become more flexible. This is particularly critical for systems related to procurement of inputs and disbursement of funds. In the area of efficiency, a necessary trade-off exists between flexibility and speed, on one hand, and accountability and transparency, on the other, when designing and implementing systems for administrative and financial project and program control.

5. UNDP faces an important challenge in its efforts to strengthen decentralized systems for evaluation, monitoring and results-based management. There is a continuing need to strengthen these systems and procedures at the decentralized program level, including regional and country programs. UNDP should focus on improving monitoring and evaluation capacity at the country office level.

Annex 1: Criteria Used to Assess Development Effectiveness

Achievement of Development Objectives and Expected Results
1.1   Programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.   1.2   Programs and projects have resulted in positive benefits for target group members.   1.3   Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.   1.4   Programs contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.

Cross Cutting Themes: Inclusive Development Which can be Sustained (Gender Equality and Environmental Sustainability)
2.1   Extent to which multilateral organization supported activities effectively addresses the crosscutting issue of gender equality.   2.2   Extent to which changes are environmentally sustainable.

Sustainability of Results/Benefits
3.1   Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term developmental results.   3.2   Projects and programs are reported as sustainable in terms of institutional and/or community capacity.   3.3   Programming contributes to strengthening the enabling environment for development.

Relevance of Interventions
4.1   Programs and projects are suited to the needs and/or priorities of the target group.   4.2   Projects and programs align with national development goals.   4.3   Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.

Efficiency
5.1   Program activities are evaluated as cost/resource efficient.   5.2   Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).   5.3   Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).

Using Evaluation and Monitoring to Improve DE
6.1   Systems and process for evaluation are effective.   6.2   Systems and processes for monitoring and reporting on program results are effective.   6.3   Results-based management systems are effective.   6.4   Evaluation is used to improve development/humanitarian effectiveness.

Annex 2: Evaluation Sample

#	Country / Commission-ing Unit	Year Published	Evaluation Title	Type	Evaluation Office	Decentra-lized
1	Lao	2011	Assessment of Development Results: Lao PDR	ADR	Yes	
2	Malawi	2011	Assessment of Development Results: Malawi	ADR	Yes	
3	Thailand	2011	Assessment of Development Results: Thailand	ADR	Yes	
4	Regional Bureau for Europe and the Commonwealth of Independent States	2011	An assessment of UNDP's contribution towards national, regional, and local public institutions' capacity	Outcome		Yes
5	Jamaica	2011	Assessment of Development Results: Jamaica	ADR	Yes	
6	El Salvador	2011	Assessment of Development Results: El Salvador	ADR	Yes	
7	Senegal	2011	Assessment of Development Results: Senegal	ADR	Yes	
8	Mongolia	2011	Assessment of Development Results: Mongolia	ADR	Yes	
9	Regional Bureau for Latin America and the Caribbean	2011	Outcome oriented mid-term RPD evaluation	Outcome		Yes
10	Ghana	2011	Assessment of Development Results: Ghana	ADR	Yes	
11	Bangladesh	2011	Assessment of Development Results: Bangladesh	ADR	Yes	
12	Nigeria	2011	Anti-corruption measures and procurement reform (PPL)	Outcome		Yes
13	Regional Bureau for Europe and the Commonwealth of Independent States	2011	Incorporation of Gender Main-streaming and Analysis into Programming in RBEC (country level/regional level)	Outcome		Yes
14	Bureau of Development Policy	2010	Evaluation of UNDP Contribution to Strengthening Local Governance	Thematic	Yes	
15	Bureau of Development Policy	2010	Evaluation of UNDP Contribution to Environmental Management for Poverty Reduction: The Poverty-Environment Nexus	Thematic	Yes	
16	Tanzania	2010	Delivering as One country-led evaluation	Other		Tanzania (Country Led)

#	Country / Commission-ing Unit	Year Published	Evaluation Title	Type	Evaluation Office	Decentra-lized
17	Operation Support Group	2010	Evaluation of UNDP Contribution at the Regional Level to Development and Corporate Results Sub-Region	Thematic	Yes	
18	Bureau of Development Policy	2010	Evaluation of UNDP Contribution to Strengthening National Capacities	Thematic	Yes	
19	Bureau for Crisis Prevention and Recovery	2010	Evaluation of UNDP Contribution to Disaster Prevention and Recovery	Thematic	Yes	
20	Somalia	2010	Assessment of Development Results: Somalia	ADR	Yes	
21	Georgia	2010	Assessment of Development Results: Georgia	ADR	Yes	
22	Cambodia	2010	Assessment of Development Results: Cambodia	ADR	Yes	
23	China	2010	Assessment of Development Results: China	ADR	Yes	
24	Bureau of Development Policy	2010	Evaluation of the Joint UNDP/World Bank/UNAIDS HIV Mainstreaming Programme	Thematic	Joint	Joint
25	Maldives	2010	Assessment of Development Results: Maldives	ADR	Yes	
26	Turkey	2010	Assessment of Development Results: Turkey	ADR	Yes	
27	Indonesia	2010	Assessment of Development Results: Indonesia	ADR	Yes	
28	Guyana	2010	Assessment of Development Results: Guyana	ADR	Yes	
29	Ethiopia	2010	Midterm Outcome Evaluation of the Governance Programme, Democratic Institution Programme (DIP)	Outcome		Yes
30	Egypt	2010	Outcome evaluation of UNDP Outcome 5: Sustainable Management of environment and natural resources incorporated into anti-poverty programming	Outcome		Yes
31	Zambia	2010	Assessment of Development Results: Zambia	ADR	Yes	
32	Myanmar	2010	Independent Assessment of the UNDP Human Initiative in Myanmar, 2010	Others		Yes
33	Sudan	2010	Review of the Strategic Partnership Framework on Governance and Rule of Law for HRBA and Gender Mainstreaming	Others		Yes
34	Regional Bureau for Europe and the Commonwealth of Independent States	2010	Midterm Assessment of the RBEC Regional Programme in Europe and CIS	Others		Yes

#	Country / Commission-ing Unit	Year Published	Evaluation Title	Type	Evaluation Office	Decentra-lized
35	Serbia	2010	Country Programme Document (2005-2009) terminal evaluation	Outcome		Yes
36	Seychelles	2009	Assessment of Development Results: Seychelles	ADR	Yes	
37	Chile	2009	Assessment of Development Results: Chile	ADR	Yes	
38	Tajikistan	2009	Assessment of Development Results: Tajikistan	ADR	Yes	
39	Uganda	2009	Assessment of Development Results Uganda	ADR	Yes	
40	Botswana	2009	Assessment of Development Results: Botswana	ADR	Yes	
41	Bosnia and Herzegovina	2009	Assessment of Development Results: Bosnia and Herzegovina	ADR	Yes	
42	Afghanistan	2009	Assessment Development Results: Afghanistan	ADR	Yes	
43	Barbados	2009	Assessment of Development Results: Barbados and OECS	ADR	Yes	
44	Kazakhstan	2009	Outcome Evaluation of UNDP Support to Democratic Governance	Outcome		Yes
45	Guatemala	2009	Assessment of Development Results: Guatemala	ADR	Yes	
46	South Korea	2009	2005-2008 CPD Terminal Evaluation	Others		South Korea
47	Sudan	2009	UNDP Sudan CCF2 (2002-2008) Evaluation	Outcome		Yes
48	Uzbekistan	2009	Assessment of Development Results: Uzbekistan	ADR	Yes	
49	Palestine	2009	Outcome Evaluation of the UNDP/PAPP Mid-Term Strategic Framework, 2008 – 2011	Outcome		Yes
50	Viet Nam	2009	HIV Programme Evaluation	Others		Yes
51	RBAS	2009	Evaluation of the third Regional Cooperation Framework for Arab States	RCF	Yes	
52	India	2009	Disaster Risk Management Outcome Evaluation	Outcome		Yes
53	Philippines	2009	Assessment of Development Results: Philippines	ADR	Yes	
54	Burkina Faso	2009	Assessment of Development Results: Burkina Faso	ADR	Yes	
55	Peru	2009	Assessment of Development Results: Peru	ADR	Yes	

Annex 3: Methodology

This Annex provides a more thorough explanation of the key elements of the methodology used for the review of the development effectiveness of UNDP. It is structured around the sequence of tasks undertaken during the review: determining the rationale for the review; drawing the sample of evaluations; undertaking the process of review and controlling for quality during the analysis phase; and, assessing the level of coverage provided by the development effectiveness review.

The review of evaluation reports was supplemented by a review of UNDP corporate documents to expand the profile of the organization and its programming. A list of the documents consulted is provided in Annex 6.

Rationale for the Development Effectiveness Review

The term "common approach" describes the use of a standard methodology, as implemented in this review, to consistently assess the development effectiveness of multilateral organizations. It offers a rapid and cost effective way to assess the development effectiveness of the organization, relative to a more time consuming and costly joint evaluation.[27] The approach was developed to fill an information gap regarding development effectiveness of multilateral organizations. Although these organizations produce annual reports to their management and/or boards, member states were not receiving a comprehensive overview of the performance on the ground of multilateral organizations. This approach complements the organizational effectiveness assessment of the Multilateral Organization Performance Assessment Network (MOPAN). The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[28]

1. There is a need for field-tested and evidence-based information on the development effectiveness of the organization.

2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-synthesis methodology to synthesize an assessment of the organization's development effectiveness.

The first condition is satisfied, as UNDP's existing reporting mechanisms do not provide sufficient information on the organization's development effectiveness. The second condition is also satisfied, as the Evaluation Office at UNDP does produce enough evaluation reports of good quality and with sufficient coverage of investments to support an assessment of the development effectiveness of UNDP. The following sections describe how the review of UNDP met these two conditions.

---

[27] "Joint evaluation" refers to a jointly funded and managed comprehensive institutional evaluation of an organization. It does not refer to DAC/UNEG Peer Reviews of the Evaluation Function.

[28] Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

UNDP's Reporting on Development Effectiveness

UNDP reports to its Executive Board on development effectiveness mainly through three regular agency-wide reporting documents: (1) the UNDP annual report, (2) the Administrator's regular report on the progress of the strategic plan, and (3), the *Annual Report on Evaluation*. These reports deal explicitly with UNDP's development performance and results.[29]

UNDP's annual report[30] highlights global activities and provides a qualitative picture of development results. This picture is further refined through detailed case examples of work in specific countries on the four strategic program areas of UNDP's mandate. Quantitative information is provided throughout the report, mainly at the output level. The report describes the extensive reach of UNDP programming as well as the number and type of beneficiaries reached. It also provides information on the portions of UNDP funding and programming going to each of the four priority program areas and to crosscutting themes such as gender equality.

The *Annual Report on Evaluation* provides clear information on the annual level of evaluation activity in all different program areas of UNDP as well as an assessment of the quality of evaluation reports. It also clearly identifies the roles of the Evaluation Office and other offices responsible for the evaluation function. Finally, the *Annual Report on Evaluation* provides an overview of the findings of evaluations carried out each year and identifies the challenges faced by UNDP in improving its efficiency and effectiveness.

The 2010 report noted positive findings around UNDP's role as a trusted and valued partner contributing to human development in the program countries. It also noted, however, challenges in the following areas:

- Making better use of "downstream" project work by better considering and planning for opportunities for "scaling up" to the national level;

- Finding the appropriate balance between short-term and long-term project investments in a given country;

- The need to better address gender issues by providing greater attention and resources and by using appropriate analysis to formulate concrete strategies;

- The need for better defined results frameworks and more concrete baselines and performance measures in project design;

- The need to develop better strategies for "graduation" and for the scaling up of pilot activities;

- The need for stronger monitoring and evaluation, knowledge management and guidance, especially for country offices; and,

- The challenge of business and operational practices, leading to delays in payments to partners or in procurement and subsequent delays in the timely completion of planned projects. This further compromises the overall efficiency and effectiveness of UNDP interventions.

Many of these challenges are also noted in the findings and conclusions of this report.

---

[29] Annual Report of the Administrator on the Strategic Plan: Performance Results for 2009, UNDP, 2010.
[30] People Centered Development: UNDP in Action - Annual Report 2010/2011, UNDP, 2011.

At a substantive level, the Administrator provides the Executive Board with regular reports on UNDP's progress in implementing the strategic plan. In 2011, there was a significant improvement in the evidence base used to report on the strategic plan. The *Midterm Review of the UNDP Strategic Plan and Annual Report of the Administrator*[31] was provided to the Executive Board in advance of its meeting in June 2011. This report makes use of the work of UNDP's Evaluation Office by making frequent references to the findings of ADRs, thematic, and outcome evaluations.

For each priority programming area, the *Midterm Review* draws on the results of a set of applicable evaluation reports (from a low of seven to a high of 13 evaluations depending on the program area). The information extracted from the evaluations includes a summary of findings for each outcome, along with examples of successes and challenges.

This approach to reporting on development effectiveness draws on UNDP's own body of published evaluative evidence as summarized in the ARE. It goes some way to meeting the need for information on development effectiveness that is backed by field-based evidence. Unfortunately, the number of evaluations used is small and findings in the *Annual Report on Evaluation* are not quantified. The *Midterm Review* does highlight some of the challenges facing UNDP in its pursuit of the goals of the strategic plan. However, it lacks a balanced reporting of the distribution of evaluation findings (both positive and negative) so that the reader could appreciate the significance of positive and negative findings.

As noted above, UNDP's own regular reporting on development effectiveness, although strengthened in 2011, lacks a balanced overview of field-tested evidence on development effectiveness. Therefore, the first condition for carrying out a development effectiveness review using the common approach is met.

UNDP's Evaluation Function (Quantity and Quality)

Quantity of UNDP Evaluations: The evaluation function at UNDP produces seven (7) different types of evaluations.[32] Section 1.3.4 in this report provides an overview of the different evaluation types and their content, organized by the organizational units responsible.

The Evaluation Resource Center website at UNDP (www.UNDP.org/Evaluations) identifies 199 evaluation reports published between the beginning of 2009 and August 2011, which corresponds to the period of the current UNDP strategic plan. This group of 199 evaluations forms the evaluation universe from which the sample of 55 evaluations was drawn for this review. The basic characteristics of both the universe (199) and the sample (55) are described further on in this Annex.

Since early 2009, UNDP's Evaluation Office has published evaluations covering programming in 32 countries and addressing specific programs or themes in a further 59. Therefore, programming in 91 countries was included in the evaluation universe (UNDP implements programming in 176 countries). Further geographic coverage is provided by evaluations focused on three of UNDP's five regions (the Arab States, Europe and the Commonwealth of Independent States, and Latin America and the Caribbean). In total, the evaluations published

---

[31] Midterm Review of the UNDP Strategic Plan and Annual Report of the Administrator, UNDP, 2011.
[32] *Evaluation Policy of UNDP,* UNDP, 2011. This does not include evaluations produced by UN Volunteers (UNV) and the United Nations Capital Development Fund.

by UNDP since early 2009 create a substantial pool of reports that is large and diverse enough to support a meta-synthesis for assessing development effectiveness.

The Quality of UNDP Evaluations: A Peer Review of the Evaluation Function at UNDP was carried out in 2005 under the auspices of the DAC Network on Development Evaluation. The review addressed the question: "Does UNDP's Evaluation Office produce evaluations which are credible, valid and useable for learning and accountability purposes as tested by internationally recognised evaluation peers?" The peer review's authors responded that:

> "The United Nations Development Programme has an Evaluation Office 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." [33]

The review team conducted its own meta-evaluation of a sample of 55 different evaluations. The results of that quality assessment were positive, with 96% of the reviewed evaluations scoring higher than 31 points out of a possible total of 45. The evaluations were scored against 11 different dimensions of quality derived from the United Nations Evaluation Group Norms and Standards for Evaluation (Annex 4).

Therefore, the second condition is met for carrying out a development effectiveness review using the common approach. There is a large enough body of evaluation reports of good quality to support the use of meta-synthesis to assess UNDP's development effectiveness.

Selecting the Evaluation Sample

The sample of evaluations was selected in three steps. First, the universe of evaluations (published between 2009 and 2011) was classified according to type (see Table 4).[34] Second, a purposive sample was drawn according to the following principles:

- Maximize geographic coverage by including a large number of ADRs from all five regions of UNDP programming and including UNDP's largest program (Afghanistan);
- Extend geographic coverage by adding a mix of thematic and outcome evaluations with a global, regional or country focus;
- Include a mixture of evaluation types; and,
- Ensure that UNDP's main program areas were adequately represented.

Third, the selected sample was compared to the universe of evaluations and examined for potential bias (Table 3).

After the sample was selected, each of the 55 evaluations was assessed in terms of the quality of the report. This sequencing was used because the process of reviewing the quality of the

---

[33] Peer Assessment of Evaluation in Multilateral Organizations: United Nations Development Programme. Ministry of Foreign Affairs of Denmark. 2005, p. 4.

[34] In addition to the 199 ADRs, thematic evaluations, outcome and other evaluations noted here, UNDP's Evaluation Resource Centre identifies 452 project evaluations completed during the 2009 to 2011 period. These were not sampled because the higher-level evaluations provided a good mix of field-tested evidence and broader global, regional and national coverage. Also, as the higher program-level evaluations encompass the initiatives at the project level, including these evaluations would be double counting the same work.

evaluations (meta-evaluation) was labour intensive, requiring almost as much time as the review of evaluation findings (the meta-synthesis). Thus, it was not practical to conduct a meta-evaluation of all evaluations available for sampling.

The sample shows an over-representation of ADRs and thematic evaluations, as compared to the larger universe of evaluations. This selection occurred deliberately, in order to provide wide coverage of as many country programs as possible. Thematic evaluations were also emphasized, as they are carried out by UNDP's Evaluation Office and cover more global programs than other evaluations. Outcome and other evaluations are somewhat under-represented in the sample, as their scope is typically quite narrow. Those thematic evaluations included in the sample were selected because they either focused on regional programs or on important programs in countries not covered by the ADRs.

The sample draws upon evaluations published during the current UNDP strategic plan. It covers programming in 44 countries (25% of the countries in which UNDP operates) and all five regions where UNDP has presence. The sample also includes evaluations on cross cutting themes, such as gender equality and environmentally sustainable development.

Most of the evaluations in the sample (37 of 55) were either country program evaluations or global thematic evaluations carried out by UNDP's Evaluation Office (i.e. centralized evaluations). Of the remaining 18 evaluations, 15 were decentralized evaluations carried out by the UNDP office responsible for the program being evaluated. Two evaluations were led by the country concerned (Tanzania and South Korea respectively) and one was a joint evaluation managed collectively by UNDP, UNAIDS and the World Bank.

Table 3: Comparison between Evaluation Universe and Sample, by Evaluation Type

Evaluation Type	Evaluation Universe (published 2009 to 2011)	Evaluation Universe (%)	Sample (published 2009 to 2011)	Sample (%)
Assessments of Development Results (ADR)	**32** ADRs covering development results in 32 different countries from all UNDP regions and including the largest single program in funding terms (Afghanistan)	16%	**31** ADRs covering results in 31 countries from all regions (Europe and the Commonwealth of Independent States, Latin America and the Caribbean, Arab States, Africa and Asia).	56%
Outcome Evaluations	**104** Outcome evaluations covering sub-national or regional programs in 44 countries and two regional bureaus (Europe and the Commonwealth of Independent States and Latin America and the Caribbean)	52%	**11** Outcome evaluations covering a further 6 countries and 1 regional evaluation of gender mainstreaming.	20%
Thematic Evaluations	**7** Thematic evaluations with global coverage of UNDP results in focus areas such as strengthening local governance and contributing to environmental management.	4 %	**6** thematic evaluations covering global programs or sectors	11%

Evaluation Type	Evaluation Universe (published 2009 to 2011)	Evaluation Universe (%)	Sample (published 2009 to 2011)	Sample (%)
Regional Cooperation Framework Evaluations	**2** Evaluations of RCFs for Arab States and for Europe and the Commonwealth of Independent Sates	1%	**1** Evaluation of RCF for Arab States	2%
Other Evaluations	A diverse mix of **50** country specific project, program and policy evaluations and program or policy evaluations at the regional level.	27%	**6** other evaluations covering 3 more countries and including the evaluation of Delivering-as-One in Tanzania.	11%

Programming Period Covered by the Evaluations in the Sample

In order to assess the period covered by the evaluations in the sample, the team reviewed each evaluation to determine its scope in terms of years of programming evaluated. All the evaluations reviewed covered at least some programming under the current UNDP strategic plan (2008-2013). As Table 4 describes, evaluation programming coverage is towards the later years. Key highlights include:

- 70% of the evaluations in the sample covered program periods ranging from 2005 to 2011.

- 10% of the evaluations reviewed covered only programming in the 2008-2011 period.

- All other evaluations also covered one or more years before the current strategic plan cycle.

- 30% of the evaluations covered two programming cycles in a single country.

- The beginning years of coverage for each of the evaluations reviewed range from 2000 to 2009.

Table 4: Programming Period Covered by the Reviewed Evaluations (*n*=55)

Beginning Year for Period Under Evaluation	Proportion of Sample[35]	Cumulative Proportion of Sample
2009 - 2011	3.5%	67.2%
2008	7.3%	
2007	16.4%	
2006	20%	
2005	20%	
2004	7.3%	27.3%
2003	5.5%	
2002	5.5%	
2001	5.5%	
2000	3.5%	
Year not specified	5.5%	5.5%
Total	**100%**	**100%**

The review team is confident that the evaluations reviewed report findings relevant to the current UNDP strategic plan period (2008-2013) given that:

- All the evaluations reviewed covered at least some programming years extending into the current strategic plan period;
- Assessment of Development Results (ADR) evaluations, while often covering two programming cycles (ranging from three to six years), concentrate their findings on the most recent program periods; and,
- Where some findings reportedly changed during the course of an evaluation period, the reviewers chose the most recent findings. For example, if an ADR reported that program changes between cycles led to more positive results in the later cycle, these results were coded as findings.

Process and Quality Assurance

The review itself was conducted by a team of five analysts, including two from CIDA's Evaluation Directorate. A two-day training session was held for analysts to build a common understanding of the review criteria. Following, the team leader and analysts conducted a pre-test in order to independently review two evaluations. The team compared their ratings for these two evaluations and developed common agreement on the classification of results for all sub-criteria. This process helped to standardize classification decisions made by the five analysts.

During the review of evaluations, analysts conferred regularly over any classification issues that arose. A second test occurred at the approximate mid-point of the review period, with all

---

[35] Where an ADR covered two program cycles, the latest cycle is used.

analysts independently rating a third evaluation. As previously, the analysts compared their respective classification decisions and resolved discrepancies.

Once the reviews were completed, the team leader reviewed the coded findings and carefully examined the cited evidence and contributing factors. Based on this examination, the team leader made a small number of adjustments to the coded findings. The process of training, testing and monitoring throughout the review process minimized any inter-analyst reliability issues and controlled for bias on the part of any one reviewer.

All 55 evaluations in the sample were found to be of sufficient quality to be included in the review. Of a possible maximum total quality score of 45, the mean score for all evaluations was 37.7. The distribution of total scores for all evaluations is described in Annex 4. Only two evaluations had a score of less than 30. The quality criterion related to evaluation findings being relevant and evidence based is of outmost importance. For this quality criterion, the mean score averaged across all evaluations in the sample was 3.9 out of a maximum score of 4.

Coverage of Effectiveness Criteria

The review team developed a set of ranges in order to assess the level of coverage of a given sub-criterion across the evaluations reviewed. Strong coverage was assigned when the number of evaluations addressing a particular sub-criterion (a) was in the range of 45 to 55. Moderate coverage for a particular sub-criterion was assigned when a was between 30 and 44. Weak coverage for a particular sub-criterion was assigned when a was less than 30.

Of the 19 sub-criteria, 13 were addressed in 45 or more evaluation reports. Therefore, these 13 sub-criteria received a rating of strong coverage. The remaining six sub-criteria received a rating of moderate coverage. No sub-criteria received a rating of weak coverage (see Table 6).

Table 5: Levels of Coverage for Each Assessment Sub-Criteria

Criteria and Sub-Criteria		*a**	Coverage **
Achievement of Development Objectives and Expected Results			
1.1	Programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.	55	High
1.2	Programs and projects have resulted in positive benefits for target group members.	47	High
1.3	Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.	43	Moderate
1.4	Programs contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	48	High
Inclusive Development which can be Sustained (Gender Equality and Environmental Sustainability)			
2.1	Extent to which multilateral organization supported activities effectively addresses the cross-cutting issue of gender equality.	50	High
2.2	Extent to which changes are environmentally sustainable.	38	Moderate
Sustainability of Results/Benefits			
3.1	Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term developmental results.	50	High
3.2	Projects and programs are reported as sustainable in terms of institutional and/or community capacity.	53	High
3.3	Programming contributes to strengthening the enabling environment for development.	49	High
Relevance of Interventions			
4.1	Programs and projects are suited to the needs and/or priorities of the target group.	53	High
4.2	Projects and programs align with national development goals.	51	High
4.3	Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations s for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	53	High
Efficiency			
5.1	Program activities are evaluated as cost/resource efficient.	39	Moderate
5.2	Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).	31	Moderate
5.3	Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).	41	Moderate
Using Evaluation and Monitoring to Improve DE			
6.1	Systems and process for evaluation are effective.	42	Moderate
6.2	Systems and processes for monitoring and reporting on program results are effective.	48	High
6.3	Results based management systems are effective.	37	Moderate
6.4	Evaluation is used to improve development/humanitarian effectiveness.	47	High
*n = number of evaluations addressing the given sub-criteria			
** Strong: *a* = 45 – 55; Moderate: *a* = 30 – 44; Weak: *a* = under 30			

Annex 4: Evaluation Quality - Scoring Guide and Results

	Evaluation Quality Scoring Guide		
	Criteria to be Scored	**Points**	**Score**
A	**Purpose of the evaluation is clearly stated.** The report describes why the evaluation was done, what triggered it (including timing in the project/program cycle) and how it was to be used.	3	
B	**The evaluation report is organized, transparently structured, clearly presented and well written.** There is a logical structure to the organization of the evaluation report. The report is well written with clear distinctions and linkages made between evidence, findings, conclusions and recommendations.	3	
C	**Evaluation objectives are stated.** Evaluation objectives are clearly presented and follow directly from the stated purpose of the evaluation.	2	
D	**Subject evaluated is clearly described.** Evaluation report describes the activity/program being evaluated, its expected achievements, how the development problem would be addressed by the activity and the implementation modalities used.	4	
E	**Scope of the evaluation is clearly defined.** The report defines the boundaries of the evaluation in terms of time period covered, implementation phase under review, geographic area, and dimensions of stakeholder involvement being examined.	4	
F	**Evaluation criteria** used to assess program effectiveness are clearly identified in the evaluation report and cover a significant number of the Common Criteria for Assessing Development Effectiveness.	5	
G	**Multiple lines of evidence are used.** The report indicates that more than one line of evidence (case studies, surveys, site visits, and key informant interviews) is used to address the main evaluation issues. One point per line of evidence to maximum of 5.	5	
H	**Evaluations are well designed.** The methods used in the evaluation are appropriate to the evaluation criteria and key issues addressed. Elements of good design include: an explicit theory of how objectives and results were to be achieved, specification of the level of results achieved (output, outcome, impact), baseline data (quantitative or qualitative) on conditions prior to program implementation, a comparison of conditions after program delivery to those before, and a qualitative or quantitative comparison of conditions among program participants and those who did not take part.	5	
I	**Evaluation findings and conclusions are relevant and evidence based.** The report includes evaluation findings relevant to the assessment criteria specified. Findings are supported by evidence resulting from the chosen methodologies. Conclusions are linked to the evaluation findings as reported.	4	
J	**Evaluation report indicates limitations of the methodology.** The report includes a section noting the limitations of the methodology. It indicates any limitations in the design as well as any problems in the implementation (low survey returns for example) and describes how their impact on the validity of results and any measures taken to address the limitations (re-surveys, follow-ups, additional case studies, etc.	5	
K	**Evaluation includes recommendations.** The evaluation report contains specific recommendations that follow on clearly from the findings and conclusions. Further, the recommendations are specifically directed to one or more organizations and are actionable and aimed at improving Development Effectiveness. (Objectives achievement, cross cutting themes, sustainability, cost efficiency or relevance).	5	
	Total Possible Score	**45**	

The criteria used for assessing evaluation quality were assigned weighted possible scores (from 2 to 5). These weights (potential total scores for a criteria) were based on the relative

importance and the complexity of a given criteria. The definition of each quality assessment criteria contains a single verifiable component, which corresponds to each available point. For example, if criteria E is worth a possible four points, then the analyst would need to verify that each of the four components of the definition were present in order to award four points for this criteria. The verified components minimized the amount of subjectivity involved in the process of quality review.

Evaluation Quality Scoring Results

During the Pilot Testing of the Common Approach, the Management Group of participating development agencies guiding the work on behalf of DAC-EVALNET suggested grouping quality score results for each evaluation into groups of five (in total score). This was seen as presenting the best level of "granularity" and transparency. It allows independent observers to reach their own conclusions on the distribution of quality scores.

Table 6: Evaluation Quality Scoring Results

Evaluation Quality Scores in Groups of Five (Max = 45)		Evaluations in Each Bracket (#)	Evaluations in Each Bracket (%)
Quality Rating Score (Groups of 5)	41-45	13	23.6%
	36-40	28	50.9%
	31-35	12	21.8%
	26-30	2	3.6%
	21-25	0	.0%
	16-20	0	.0%
	11-15	0	.0%
	6-10	0	.0%
	0-5	0	.0%
	Total	**55**	**100.0%**

Given the distribution shown in Table 6, it would have been possible to establish a cut-off score of 30 and to exclude the two evaluation reports scoring in the 26-30 range. However, on close examination, these two evaluation reports had high scores on core criteria (especially criterion I, "evaluation's findings and conclusions are relevant and evidence based") and so were included in the analysis.

While the overall quality of reviewed evaluations is quite high, UNDP's Evaluation Office is continuously improving its practices. Thus, there may be a difference in quality or rigour between recently published evaluations and older ones. Variation in quality is one reason why the evaluation sample was restricted to a relatively short publication period (January 2009 to September 2011). The meta-evaluation results indicate that, although there may be some improvements in evaluation quality over the course of the review period, all of the sampled evaluations met the review's quality standards.

Annex 5: Guide for Classifying Evaluation Findings

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
Common Development Evaluation Assessment Criteria				
1. Achievement of Development Objectives and Expected Results				
1.1 MO supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.	Evaluation finds that Less than half of stated output and outcome objectives have been achieved including one or more very important output and/or outcome level objectives.	Evaluation finds that half or less than half of stated objectives (at the output and outcome level) are achieved.	Evaluation finds that MO supported programs and projects either achieve at least a majority of stated output and outcome objectives (more than 50% if stated) or that the most important of stated output and outcome objectives are achieved.	Evaluation finds that MO supported programs and projects achieve all or almost all significant development and/or humanitarian objectives at the output and outcome level.
1.2 MO supported programs and projects have resulted in positive benefits for target group members.	Evaluation finds that problems in the design or delivery of MO supported activities mean that expected positive benefits for target group members have not occurred or are unlikely to occur.	Evaluation finds that that MO supported projects and programs result in no or very few positive changes experienced by target group members. These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.	Evaluation finds that MO supported projects and programs have resulted in positive changes experienced by target group members (at the individual, household or community level). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.	Evaluation finds that MO supported projects and programs have resulted in widespread and significant positive changes experienced by target group members as measured using either quantitative or qualitative methods (possibly including comparison of impacts with non-program participants). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.
1.3 MO programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.	Evaluation finds that MO supported projects and programs have not contributed to positive changes in the lives of beneficiaries as measured quantitatively or qualitatively.	Evaluation finds that MO supported projects and programs have contributed to positive changes in the lives of only a small number of beneficiaries (when compared to project or program targets and local or national goals if established).	Evaluation finds that MO supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries as measured quantitatively or qualitatively. These may result from development, relief, or protracted relief and rehabilitation operations and may include the avoidance of negative effects of emergencies.	Evaluation finds that MO supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries. Further, they have contributed to the achievement of specific national development goals or have contributed to meeting humanitarian relief objectives agreed to with the national government and/or national and international development and relief organizations.
1.4 MO activities contributed to significant changes in national	Evaluation finds that national policies and programs in a given sector or area of	Evaluation finds that MO activities have not made a significant contribution to the	Evaluation finds that MO activities have made a substantial contribution to either	Evaluation finds that MO activities have made a substantial contribution to either

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.	development (including disaster preparedness, emergency response and rehabilitation) were deficient and required strengthening but MO activities have not addressed these deficiencies.	development of national policies and programs in a given sector or area of development, disaster preparedness, emergency response or rehabilitation. (Policy changes in humanitarian situations may include allowing access to the effected populations).	re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation.	re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation. Further, the supported policies and program implementation modalities are expected to result in improved positive impacts for target group members.
2. Cross Cutting Themes: Inclusive Development Which can be Sustained				
2.1 Extent MO supported activities effectively address the cross-cutting issue of gender equality.	Evaluation finds that MO supported activities are unlikely to contribute to gender equity or may in fact lead to increases in gender inequities.	Evaluation finds that MO supported activities either lack gender equality objectives or achieve less than half of their stated gender equality objectives. (Note: where a program or activity is clearly gender focused (maternal health programming for example) achievement of more than half its stated objectives warrants a satisfactory rating).	Evaluation finds that MO supported programs and projects achieve a majority (more than 50%) of their stated gender equality objectives.	Evaluation finds that MO supported programs and projects achieve all or nearly all of their stated gender equality objectives.
2.2 Extent changes are environmentally sustainable.	Evaluation finds that MO supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. In addition the evaluation reports that changes resulting from MO supported programs and projects are not environmentally sustainable.	Evaluation finds that MO supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. There is, however, no direct indication that project or program results are not environmentally sustainable. OR Evaluation finds that MO supported programs and projects include planned activities or project design criteria intended to promote sustainability but these have not been successful.	Evaluation finds that MO supported programs and projects include some planned activities and project design criteria to ensure environmental sustainability. These activities are implemented successfully and the evaluation reports that the results are environmentally sustainable.	Evaluation finds that MO supported programs and projects are specifically designed to be environmentally sustainable and include substantial planned activities and project design criteria to ensure environmental sustainability. These plans are implemented successfully and the evaluation reports that the results are environmentally sustainable.
3. Sustainability of Results/Benefits				
3.1 Benefits continuing or likely to	Evaluation finds that there is a very low	Evaluation finds that there is a low	Evaluation finds that it is likely that the	Evaluation finds that it is highly likely that

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term developmental results.	probability that the program/project will result in continued intended benefits for the target group after project completion. For humanitarian relief operations, the evaluation finds no strategic or operational measures to link relief, to rehabilitation, reconstruction and, eventually, to development.	probability that the program/project will result in continued benefits for the target group after completion. For humanitarian relief operations, efforts to link the relief phase to rehabilitation, reconstruction and, eventually, to development are inadequate. (Note, in some circumstances such linkage may not be possible due to the context of the emergency. If this is stated in the evaluation a rating of satisfactory can be given)	program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the evaluation finds credible strategic and operational measures linking relief to rehabilitation, reconstruction and, eventually, development.	the program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the evaluation finds credible strategic and operational measures linking relief to rehabilitation, reconstruction and, eventually, development. Further, they are likely to succeed in securing continuing benefits for target group members.
3.2 Extent MO supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.	Evaluation finds that the design of MO supported programs and projects failed to address the need to strengthen institutional and/or community capacity as required. In the case of humanitarian operations, the design of programs and projects failed to take account of identified needs to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.	Evaluation finds that MO programs and projects may have failed to contribute to strengthening institutional and/or community capacity or, where appropriate, to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.	Evaluation finds that MO programs and projects may have contributed to strengthening institutional and/or community capacity but with limited success	Evaluation finds that either MO programs and projects have contributed to significantly strengthen institutional and/or community capacity as required or institutional partners and communities already had the required capacity to sustain program results.
3.3 Extent MO development programming contributes to strengthening the enabling environment for development.	Evaluation finds that, for development programs, there were important weaknesses in the enabling environment for development (the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting	Evaluation finds that MO development activities and/or MO supported projects and programs have **not made a notable** contribution to changes in the enabling environment for development.	Evaluation finds that MO development activities and/or MO supported projects and programs have made a **notable** contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures,	Evaluation finds that MO development activities and/or MO supported projects and programs have made a **significant** contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures,

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	structures such as capital and labor markets). Further, the MO activities and support provided to programs and projects failed to address the identified weakness successfully, further limiting program results.		service delivery and quality; and necessary improvements to supporting structures such as capital and labor markets.	service delivery and quality; and necessary improvements to supporting structures such as capital and labor markets. Further, the evaluation reports that these improvements in the enabling environment are leading to improved development outcomes.
4. Relevance of Interventions				
4.1 MO supported programs and projects are suited to the needs and/or priorities of the target group	Evaluation finds that substantial elements of program or project activities and outputs were unsuited to the needs and priorities of the target group.	Evaluation finds that no systematic analysis of target group needs and priorities took place during the design phase of developmental or relief and rehabilitation programming or the evaluation report indicates some evident mismatch between program and project activities and outputs and the needs and priorities of the target group.	Evaluation finds that the MO supported activity, program or project is designed taking into account the needs of the target group as identified through a process of situation or problem analysis (including needs assessment for relief operations) and that the resulting activities are designed to meet the needs of the target group.	Evaluation identifies methods used in program and project development (including needs assessment for relief operations) to identify target group needs and priorities (including consultations with target group members) and finds that the program and project takes those needs into account and is designed to meet those needs and priorities (whether or not it does so successfully).
4.2 MO supported development projects and programs align with national development goals:	Evaluation finds that significant elements of MO supported development program and project activity run counter to national development priorities with a resulting loss of effectiveness.	Evaluation finds that a **significant portion** (1/4 or more) of the MO supported development programs and projects subject to the evaluation are not aligned with national plans and priorities but there is no evidence that they run counter to those priorities.	Evaluation finds that **most** MO supported development programs and projects are fully aligned with national plans and priorities as expressed in national poverty eradication and sector plans and priorities. Wherever MO supported programs and projects are reported in the evaluation as not directly supportive of national plans and priorities they do not run counter to those priorities.	Evaluation finds that **all** MO supported development projects and programs subject to the evaluation are fully aligned to national development goals as described in national and sector plans and priorities, especially including the national poverty eradication strategy and sector strategic priorities.
4.3 MO has developed an effective partnership with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations for planning, coordination and	Evaluation finds that the MO experiences significant divergence in priorities from those of its (government, non-governmental organization or donor) partners and lacks a strategy or plan which will credibly address the divergence and which should result in strengthened	Evaluation finds that the MO has experienced significant difficulties in developing an effective relationship with partners and that there has been significant divergence in the priorities of the MO and its partners.	Evaluation finds that the MO has improved the effectiveness of its partnership relationship with partners over time during the evaluation period and that this partnership was effective at the time of the evaluation or was demonstrably improved.	Evaluation finds that the MO has consistently achieved a high level of partnership during the evaluation period.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.	partnership over time.			
5. Efficiency				
5.1 Program activities are evaluated as cost/resource efficient:	Evaluation finds that there is credible information indicating that MO supported programs and projects (development, emergency preparedness, relief and rehabilitation) are not cost/resource efficient.	Evaluation finds that the MO supported programs and projects under evaluation (development, emergency preparedness, relief and rehabilitation) do not have credible, reliable information on the costs of activities and inputs and therefore the evaluation is not able to report on cost/resource efficiency. OR Evaluation finds that MO supported programs and projects under evaluation present mixed findings on the cost/resource efficiency of the inputs.	Evaluation finds that the level of program outputs achieved (development, emergency preparedness, relief and rehabilitation) when compared to the cost of program activities and inputs are appropriate even when the program design process did not directly consider alternative program delivery methods and their associated costs.	Evaluation finds that MO supported (development, emergency preparedness, relief and rehabilitation) programs and projects are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time.
5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)	Evaluation finds that less than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time, there is no credible plan or legitimate explanation found by the evaluation that would suggest significant improvement in on-time objectives achievement in the future.	Evaluation finds that less than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time but the program or project design has been adjusted to take account of difficulties encountered and can be expected to improve the pace of objectives achievement in the future. In the case of humanitarian programming, there was a legitimate explanation for the delays.	Evaluation finds that more than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time and that this level is appropriate to the context faced by the program during implementation, particularly for humanitarian programming.	Evaluation finds that nearly all stated output and outcome level objectives of MO supported programs and projects are achieved on time or, in the case of humanitarian programming, a legitimate explanation for delays in the achievement of some outputs/outcomes is provided.
5.3 Evaluation indicates that MO systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical	Evaluation finds that there are serious deficiencies in agency systems and procedures for project/program implementation that result in significant delays in project start-up, implementation or completion and/or significant cost increases.	Evaluation finds some deficiencies in agency systems and procedures for project/program implementation but does not indicate that these have contributed to delays in achieving project/program objectives.	Evaluation finds that agency systems and procedures for project implementation are reasonably efficient and have not resulted in significant delays or increased costs.	The evaluation reports that the efficiency of agency systems and procedures for project implementation represent an important organizational strength in the implementation of the program under evaluation.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
arrangements etc.)				
6. Using Evaluation and Monitoring to Improve Development Effectiveness				
6.1 Systems and process for evaluation are effective.	Evaluation specifically notes that evaluation practices in use for programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are seriously deficient.	Evaluation finds that there is no indication that programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are subject to systematic and regular evaluations.	Evaluation finds that program being evaluated is subject to systematic and regular evaluations or describes significant elements of such practice. No mention of policy and practice regarding similar programs and projects. This may include specialized evaluation methods and approaches to emergency preparedness, relief and rehabilitation programming.	Evaluation finds that program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.
6.2 Systems and processes for monitoring and reporting on program results are effective	Evaluation finds an absence of monitoring and reporting systems for the development and humanitarian programming. This would include the absence of adequate monitoring of outputs during the implementation of humanitarian programming.	Evaluation finds that while monitoring and reporting systems for the development and humanitarian programming exist, they either do not report on a regular basis or they are inadequate in frequency, coverage or reliability.	Evaluation finds that monitoring and reporting systems for development and humanitarian programming as appropriate are well established and report regularly.	Evaluation finds that monitoring and reporting systems for the program are well established and report regularly. The quality of regular reports is rated highly by the evaluation and results are reportedly used in the management of the program.
6.3 Results-Based Management (RBM) systems are effective	Evaluation finds that there is no evidence of the existence of an RBM system for the program and no system is being developed.	Evaluation finds that while an RBM system is in place, or being developed, it is unreliable and does not produce regular reports on program performance.	Evaluation finds that an RBM system is in place and produces regular reports on program performance.	Evaluation finds that an RBM system is in place for the program and there is evidence noted in the evaluation that the system is used to make changes in the program to improve effectiveness.
6.4 MO makes use of evaluation to improve development/ humanitarian effectiveness	Evaluation report does not include a management response and does not have one appended to it or associated with it. There is no indication of how the evaluation results will be used. There is no indication that similar evaluations have been used to improve effectiveness in the past.	Evaluation report includes a management response (or has one attached or associated with it) but it does not indicate which recommendations have been accepted: or there is some, non-specific indication that similar evaluations have been used to improve program effectiveness in the past.	Evaluation reports includes a management response (or has one attached or associated with it) that indicates which recommendations have been accepted. Or there is a clear indication that similar evaluations in the past have been used to make clearly identified improvements in program effectiveness.	Evaluation reports includes a management response (or has one attached or associated with it) describes a response to each major recommendation which is appropriate and likely to result in the organizational and programmatic changes needed to achieve their intent.

Annex 6: Corporate Documents Reviewed

DAC. *Peer Assessment of Evaluation in Multilateral Organizations: United Nations Development Programme.* Ministry of Foreign Affairs of Denmark. 2005.

MOPAN. *MOPAN Common Approach: UNDP.* 2010.

UNDP. *Annual report of the Administrator on the Strategic Plan Performance and Results for 2009.* 2010.

UNDP. *Annual Report on Evaluation.* 2010.

UNDP. *DRAFT Report of the Second Regular Session 2011.* Executive Board of UNDP (DP/2012/1). 2012

UNDP. *Evaluation Policy of UNDP*, 2011.

UNDP. *Midterm Review of the UNDP Strategic Plan and Annual Report of the Administrator.* 2011.

UNDP. *People Centered Development: UNDP in Action - Annual Report 2010/2011.* 2011.

UNDP. *Report of the Annual Session: Executive Board of UNDP (DP/2011/31).* 2011.

UNDP. *UNDP Strategic Plan, 2008-2013: Accelerating Global Progress on Human Development.* 2008.

Annex 7: Recent Changes at UNDP

Following the presentation of this review to UNDP staff, they indicated that the organization had made a number of changes to address issues raised in its evaluations conducted between 2009 and 2011. They provided this input to be included as an annex to the report.

UNDP Comments on CIDA Report: Development Effectiveness Review of UNDP (2005-2011)

UNDP notes that the draft report recognizes and appreciates the complex political and developmental context that UNDP operates in, and which in turn informs the scope and depth of UNDP's work at all levels. UNDP welcomes the report's finding that UNDP's mandate is aligned with CIDA's thematic priorities of stimulating economic growth and increasing food security.

UNDP commends Canada for leading the way among UNDP's donor partners in optimizing resources and significantly reducing transaction costs of reviews by utilizing published and publicly available UNDP evaluation reports that encapsulate evidence-based progress being made and lessons learned at country, regional and global levels. UNDP further recognizes that the methodology that CIDA has adopted is a best practice and calls for its further emulation by UNDP's other donor partners going forward.

Both UNDP and CIDA acknowledge, however that a meta-synthesis of previously published evaluations is by nature, retrospective, and hence unable to fully capture ongoing initiatives/measures that have been undertaken (several during the latter part of the evaluation period under review) in order to appropriately address issues identified by the recommendations/findings of previous evaluations. Details of some of these measures will be provided by way of an explanatory note, below, in order to comprehensively describe UNDP's responsiveness to issues and challenges being currently faced.

Initiatives undertaken by UNDP to address issues identified by evaluations from 2010-till date:

- **Programme/project Planning and Monitoring:**
  - ✓ In response to the Mid-Term Review (MTR) of the UNDP Strategic Plan 2008-2013, UNDP has revised its development results framework, framing a sharper set of corporate outcomes.

  - ✓ The revised development results framework has included since 2011, new SMART indicators for in-depth analysis of outcomes, as well as development effectiveness and output indicators that allow for monitoring and reporting of UNDP's cooperation in response to recurring evaluation findings, and allow for disaggregation by type of contributions provided at country level.

  - ✓ The organization has also introduced a new 'Annual Business Plan' (ABP) in 2011, which aligns and focuses UNDP response to demand for programmatic interventions at global, regional and country levels pertaining to those critical development

priorities that are most likely to trigger a positive, multiplier effect and impact, within the scope of UNDP's development results framework. The ABP and associated integrated work plans also include a series of indicators to track progress on substantive and managerial results that are identified as key priorities.

✓ In 2010-11, UNDP commissioned an organization wide structural review, which led to the articulation and implementation in 2011, of an 'Agenda for Organisational Change' (AOC), which seeks to further improve the organisation's ability to deliver concrete development results at the country level, and whose three pillars include – (1) improved governance, (2) strengthened organizational effectiveness, and (3) enhanced leadership, culture, and behaviours.

✓ Looking ahead, UNDP, since 2011 is making preparations for the next Strategic Plan including scenario planning that allows for designing a more robust results framework and logic results chain, including indicators for outcome and output level contributions.

✓ UNDP's periodic corporate performance monitoring mechanism - the 'Country Office Scans' – was strengthened in 2010 to provide a solid evidence base for a substantive dialogue between senior management at HQ and country offices on programme and operations performance.

✓ In 2011, UNDP conducted substantive analysis of all 2,791 outcome indicators that are derived from Executive Board-approved programme documents (as contained in the Results and Resources Frameworks of global programmes, regional programmes and country programmes), with a view to ascertain if they are SMART, and whether they are being used properly for substantive performance management. As a result of this action, UNDP has:

o Strengthened its indicator usage by introducing an indicator in the corporate 'Balance Score Card' to rate the quality of the usage of indicators used for monitoring progress towards achievement of development results, in terms of how systematically and frequently they are used;

o Strengthened country programme formulation, including more support and training at the United Nations Development Assistance Framework roll-out stage, and a more robust appraisal process prior to submission of programme documents to UNDP's Executive Board.

o Revised existing country programme document (CPD) guidance to strengthen appraisal of indicators contained in programming documents' Results and Resources Frameworks, with an emphasis on high quality gender indicators and data.

o UNDP has also stepped up training on RBM issues including the selection and usage of SMART indicators, baselines and targets, in consonance with national systems, and has conducted training of all staff in this regard.

- **Programme/Project Reporting and Communications:**

  ✓ A complete redesigning was conducted in 2011 of the template used to record information that is provided annually by each programme unit to senior management on programmatic results, and which allows for a clear demonstration of results achieved and also underpins the Annual Report of the Administrator to the Executive Board. This redesigned template is now capable of providing a much more robust view of programmatic performance per unit, regionally and in different country settings. This template now includes specific questions on 'value for money', which is also set to help develop a baseline for improved results in this area.

  ✓ In 2011, UNDP committed to better communicating on development results, and as an important element of UNDP's 'Agenda for Change' (AOC), the UNDP corporate website was redesigned to facilitate improved access by all stakeholders and interested parties in gaining relevant information on UNDP activities. UNDP has also finalized the construction of a new website, which provides open access to information on project-level performance within the scope of the International Aid Transparency Initiative (IATI) that UNDP adheres to.

- **"Fast Track Procedures" to improve efficiency / effectiveness in crisis settings:**

  ✓ UNDP Fast Track Policies and Procedures (FTP) were created to provide increased operational flexibility to UNDP Country Offices in times of crisis and operating in special development situations without compromising accountability. The goal is to speed up UNDP's delivery of programmes, projects, or services by significantly reducing the time it takes to carry out some of the administrative and operational key procedures in the areas of procurement, finance and human resources. Decreasing the amount of time an Office has to spend on buying goods and services, recruiting staff and consultants, or making financial payments ensures that the organization can focus its recovery efforts as soon as possible after a crisis.

  ✓ FTPs have been officially adopted by UNDP in March 2010 but had already been used successfully during the Haiti earthquake in January 2010. To date around 30 Country Offices operating in crises such as Pakistan, Yemen, and Libya have used fast track procedures to allow for a more efficient delivery of their programmes and services.

  ✓ Upon request by a Country Office and triggered by an in-country crisis that can be sudden or prolonged FTPs get approved by UNDP HQ and can be activated for up to 12 months.

- ✓ All new country programme documents (CPDs) that have been developed since 2010 include mandatory provisions that allow for utilization of "Fast Track procedures" in case of crisis.

- **Efficiency issues:**

  - ✓ Within the scope of the Agenda for Organizational Change (AOC), in 2011, UNDP rolled out a "Streamlining Recruitment" project, which offers a set of concrete options for enhancing the efficiency of recruitment processes by reducing the timeframes by at least 6 weeks, in recruiting persons on international Fixed Term Appointments (FTA), and simplifying and shortening processes while maintaining full transparency and accountability.

  - ✓ A further measure for enhancing the transparency and accountability of the organization included the alignment/compliance with and rollout in 2011, of the International Public Sector Accounting Standard (IPSAS).

  - ✓ Since 2010, UNDP has been implementing several tactical projects under a 'Procurement Roadmap' within the Agenda for Organisational Change in order to further strengthen and streamline its procurement and realise more value for money. Under these projects UNDP's Bureau of Management has introduced important changes to the way business units and country offices carry out procurement. These include:

    - o Creating a 'Single Layer Review', which is a streamlining project that reduces 6 weeks off procurement processes, including recruitment of staff;

    - o Raising the threshold for small-value purchases to 5,000 USD and introducing 'Micro-Purchasing', which is a quicker and simpler method compared to the RFQ system that allowed canvassing over the phone, internet, and in person;

    - o Setting up a new system for granting Increased Delegation of Procurement Authority (IDPA), based on demand, capacity, and known risks. The system provides country offices (CO) with increased authority and reduce overall procurement lead times when certain performance criteria are met. It also informs HQ how to best support CO procurement activities;

    - o Reformulating the policy on Individual Contracts (IC) to make the modality more flexible. Oversight has been streamlined within the IDPA, and individuals may be engaged to support certain staff functions for up to 3 months, and the approval of contracts exceeding 12 months may be delegated to a Desk Officer. This ensures that ICs are used effectively and efficiently throughout the organization.