The evolution of independent country programme evaluation in UNDP
The Independent Evaluation Office (IEO) of the United Nations Development Programme (UNDP) publishes the Occasional Papers series on topics in the field of evaluation related to IEO work, including *inter alia*: Assessments of Development Results (ADRs); thematic evaluations; global or regional programme evaluations; decentralized evaluations; impact evaluations; evaluations commissioned by programme units; evaluation methods and guidance; and national evaluation capacities development. They are intended primarily for those concerned with these matters in international organizations, government, civil society and the academic community.

**Authors:**

Michael Reynolds, Independent Development Consultant
M. Nurul Alam, Independent Consultant
Michael F. Craft, Evaluation Specialist, UNDP Independent Evaluation Office

Copyright © 2016
by the United Nations Development Programme
1 UN Plaza, New York, NY 10017, USA

The views expressed in this publication are those of the author(s) and do not necessarily represent those of the United Nations, including UNDP, or the UN Member States.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission. To seek permission please email ieo@undp.org.
# Table of Contents

I. Introduction .................................................................................................................. 1  
   1.1 The Assessment of Development Results .......................................................... 1  
   1.2 Overview of the paper ......................................................................................... 2  
II. UNDP evaluation architecture .................................................................................. 2  
   2.1 Evaluation in UNDP’s first 30 years ................................................................. 2  
   2.2 The dawn of RBM .............................................................................................. 4  
   2.3 The development of the UNDP evaluation policy ............................................. 6  
III. Country programme evaluation before the ADR ...................................................... 7  
   3.1 Early efforts at country level review .................................................................. 7  
   3.2 Experimenting with country evaluation ............................................................ 8  
   3.3 Piloting the Country Level Impact Assessment ............................................... 9  
IV. Rolling out the first generation (2002-2007) ......................................................... 10  
   4.1 Original design and guidance .......................................................................... 10  
   4.2 Scope and methodology ................................................................................. 11  
   4.3 Review and reform ......................................................................................... 13  
V. Acting on a policy mandate (2008-2015) ................................................................. 13  
   5.1 Scaling report volume ..................................................................................... 14  
   5.2 Consolidation of methodology ....................................................................... 15  
   5.3 Emphasising national ownership ................................................................... 15  
   5.4 Changing nature of the consultants ............................................................... 16  
   5.5 Enhancing corporate alignment and use .......................................................... 17  
VI. ‘Evaluating as one’ in the UN system ................................................................... 18  
   6.1 Country programme evaluation behind UN reform ....................................... 21  
   6.2 ADRs in an integrated UN development system .............................................. 19  
   6.3 The spirit of increased coverage .................................................................... 20  
VII. Conclusions and lessons ....................................................................................... 22  

Annexes ............................................................................................................................ 23  
   Annex 1. Documents Consulted ............................................................................ 24  
   Annex 2. Strategies for increased ADR coverage ............................................... 26  
   Annex 3. List of ADRs ............................................................................................ 27
Abstract

Independent country programme evaluation is a central means for assessing performance and advancing learning in a highly decentralized organization, such as the United Nations Development Programme (UNDP). Assessments of Development Results (ADRs) were launched in 2002 by UNDP as a corporate tool for evaluating development results and enhancing substantive accountability at the country level. They present distinct methodological challenges for assessing results at the aggregate level in an organization with a wide scope of activities, and have undergone continuous reform since their inception. These changes occurred within the framework of an independent evaluation function in UNDP, which was firmly established by the late 1990s. Lessons from the evolution of ADRs are relevant for wider application to country programme evaluations of UN partners.

This paper seeks to provide a detailed history of ADRs after the production of 90 reports covering more than half of UNDP’s country and multi-country programmes. It begins with a background of the evaluation function in UNDP, notably the organizational policies that provided a methodological footing for the emergence of ADRs. These included the work of centralized evaluation units and a number of pilot initiatives that used the country programme as the unit of analysis. Next, the paper explores the principles, approaches and production of ADRs over two distinct phases through its current institutionalization. An iterative process of review and reform has contributed to enhancements in formalizing the policy mandate and methodology of ADRs. Finally, the paper frames UNDP’s work within the wider context of country programme evaluation in the UN Development System, both from the perspective of UN agency-specific and joint evaluations. It argues that if the evaluation function is not to continue lagging behind wider UN reform at the country level, UN partners should examine how to better integrate overlapping evaluation exercises. It concludes by suggesting areas where joint evaluations may learn from the experience of ADRs.
I. Introduction

For UNDP, a country programme evaluation (CPE) is an assessment of the entirety of its work in a particular programme country. Such evaluations are crucial in highly decentralized organisations that promote national ownership of programmes and activities. They are carried out regularly by a number of international organizations. For UNDP, nearly 90 percent of programme funds are spent at the country level where the organization arguably makes the greatest contribution to its corporate development goals. Recognizing these facts, the Evaluation of the UNDP Strategic Plan 2008-2013 concluded that:

[UNDP’s] performance monitoring and reporting system...is not optimal for a highly decentralized organization working on complex development issues where context is extremely important. This is an issue at the heart of its performance monitoring and reporting challenges. To create a more appropriate performance monitoring and assessment system, the country programme should be the unit of analysis. (UNDP 2013a)

Independent CPEs have an important role to play in supporting the idea of the country programme as the unit of analysis. They are however not without distinct challenges. First, the wide scope of a development agency’s work at the country level typically requires a broad range of skills in members of the evaluation team, as well as methodological approaches related to synthesis. Second, as an aggregation of results of many diverse activities that may not follow the same timeline, CPEs must adopt an approach that goes beyond project evaluation. In contrast to much development evaluation focused on individual projects and programmes, CPEs aggregate a higher level of results. They build on project and programme evaluations where they exist and meet adequate standards of impartiality and quality. What these aggregated evaluations have in common is the likelihood that the overall findings will be less precise than those identified in narrower project evaluations.

These methodological issues inherent to CPEs in the UNDP context are heightened further in UN joint country evaluations, such as evaluations of UN Development Action Frameworks (UNDAFs) or Delivery as One (Dao) initiatives. The dual challenge of an aggregation of results and across a large range of areas in such joint evaluations is made greater still by the number of UN partners included in the evaluation’s scope.

1.1 The Assessment of Development Results

In 2002, UNDP launched a series of independent CPEs using a new methodology called Assessment of Development Results (ADRs). The ADRs are independent evaluations that assess UNDP’s contribution to development results within a specific programme country. They are part of UNDP’s evaluation system, which aims to promote quality assurance, strengthen learning within UNDP, and support country level programming.

The ADR was developed in a rapidly moving UNDP context and was a response to the introduction of a new results-based management (RBM) system in the organization. Moreover, its introduction coincided with the changing orientation of UNDP’s work towards more upstream interventions and a period of growing resources for the organization. UNDP’s Executive Board increasingly demanded more accountability for results at the country and the organizational level. The organization was in need of a credible and evidence-based mechanism with which to respond to that demand.

The ADR was conceptualized by the UNDP Evaluation Office (EO) in response to this emerging demand by the governing body. As the first generation of ADRs were rolled out and presented to its Executive Board, UNDP introduced a new evaluation policy in 2006. The policy provided the EO with a corporate mandate and contributed to the institutional acceptance of the ADR as a corporate tool for assessing development effectiveness and substantive

---

1 For example, country portfolio or programme evaluations carried out by IFAD, UNFPA, GEF, FAO, ILO, UNIDO, and WFP.

2 The Evaluation Office (EO) was renamed the Independent Evaluation Office (IEO) in January 2014, and this paper accordingly uses IEO in reference to events occurring from 2014 onward.
accountability of its programmes at the country level.

This study of the evolution of the ADR comes after 14 years, during which 90 reports were produced. They covered more than half of UNDP’s country and multi-country programmes (Annex 3). The methodology and process used for ADRs has not remained constant. ADRs began with an exploratory approach combining the experience of CPEs and mid-term reviews. The methodology accumulated its robustness through regular reviews and reform which infused new dimensions of analysis and innovative methodologies. Through multiple external workshops and other fora, feedback was provided by governments, UNDP country offices (COs) and evaluation professionals on the relevance, process and utility of the ADR. ADRs were also subject to constant reflection by EO managers with almost annual review of both processes and methodologies.

II. UNDP evaluation architecture

Before looking at how CPE evolved, this section provides an overview of the evolution of UNDP’s evaluation architecture. It begins with a brief discussion of the first 30 years of evaluation policy and practice (1966-1996) before covering comprehensive reforms that started in 1997.

2.1 Evaluation in UNDP’s first 30 years

When UNDP was established in 1966, evaluation tended to focus more on problems of project implementation than on outputs and the achievement of objectives (UNDP, 1983: p.8). This followed the tradition established by its predecessor organizations, the Special Fund and the Extended Programme of Technical Assistance. In 1967, an independent Evaluation Division of seven staff was established. Its job was “to carry out a programme of sectoral, project, country programme and special evaluation studies and report on their results to the Administrator and Governing Council” (Joint Inspection Unit 1983: p.2). Only a few years later, the issue of evaluation was included in the comprehensive study of the capacity of the UN Development System – the ‘Capacity Study’ or Jackson Report (United Nations, 1969). The report, referring to the whole UN Development System, noted that:

Quantitatively, so much evaluation is now being attempted that it almost amounts to international hypochondria. It is a definite brake on the capacity of the system. Qualitatively, the position is the more disturbing, for very few people have the necessary experience and understanding to undertake this exacting function successfully (p.11).

This lack of an orderly and coordinated approach led the Capacity Study to recommend strongly that UNDP develop more effective evaluation machinery, both for built-in operational control of projects and

1.2 Overview of the paper

This occasional paper was originally prepared as background for ADR reform as a way of systematically placing its history on record. It has been edited for a wider audience within the UN and the evaluation community interested in CPE. First, it introduces the overall evaluation context in UNDP within which ADRs function. It then explores the evolution of CPE before discussing ADR implementation over two periods: its initial exploratory stage (2002-2007), and the current period following a revised evaluation policy (2008-2015). The final sections examine key issues that UNDP will continue to face while producing ADRs, namely their role in the context of the UN Development System’s evolving joint evaluation at the country level.

The paper used a wide range of resources including Executive Board reports, decisions, and papers prepared by UNDP. The analysis of the first phase of ADR implementation drew on peer reviews of UNDP’s evaluation function and the various guidelines developed for the ADR and its predecessors.

---

3 The 48-member Governing Council was superseded by the Executive Board on 1 January 1994.
programmes and for a systematic analysis and assessment of the results achieved. The Capacity Study proposed a comprehensive, controlled evaluation system as part of a decentralized management system. Evaluation was to be an indispensable and “constant thread running through all other phases of the development cooperation cycle” (Joint Inspection Unit, 1983: p.2). The Capacity Study further proposed the creation of a special inspection and evaluation unit at headquarters that would have overall responsibility for the assessment of project and programme results. This particular recommendation was not incorporated into the agreement on implementing reform of UNDP – the Consensus – approved by the Governing Council in 1970. On the contrary, evaluation did not have a high profile in the Consensus (UNDP 1983).

In 1973, however, the Evaluation Division was disbanded and merged into the Programme Planning and Policy staff in UNDP headquarters and then to the Bureau for Programme Policy and Evaluation (BPPE) when it was established in 1976. What began in UNDP as a “comprehensive and orderly evaluation system development effort during the late 1960s and early 1970s somehow lost its momentum during the mid-1970s”, which led to concerns over ensuring the effectiveness of the evaluation system (Joint Inspection Unit 1983: para.8).

Individual project evaluation responsibility in UNDP was decentralized to the regional bureaux (RBx), field offices, executing agencies and governments. Thematic evaluations were the responsibility of BPPE with only one person serving part-time as senior evaluation coordinator and other BPPE technical advisers taking an active part in studies in their respective fields of competence. However, BPPE responsibilities for overall evaluation system development, oversight, training, and follow-up were unclear, and the RBx evaluation approaches varied from a formal unit with emphasis on design to little specific activity (Joint Inspection Unit, 1981). During this period, evaluation as an organized part of the project cycle management was marginalized in the practices of UNDP’s field and headquarter units. Periodic project reviews were considered more appropriate and were accepted as valid basis for management decisions on projects and programmes. An initiative called the ‘special project implementation review exercise’ (SPIRE) was introduced in 1982 by UNDP’s headquarters. This aimed to developed a streamlined approach for project monitoring and review. In view of a lack of any specific corporate policy and prescribed methodology for project or programme level evaluation for the COs, SPIRE came closest to what may be called an evaluative approach for project assessment.

Recognising increasing concern about the effectiveness and impact of development cooperation, and specifically the Governing Council’s wish to be better informed about the results of UNDP’s activities, the UNDP Administrator presented an analysis of the evaluation function to the council at the second Session in 1983 (UNDP 1983). The Council noted “with satisfaction the decision of the Administrator to establish a small central evaluation unit” (Decision 83/12). The Central Evaluation Office (CEO) was established in October 1983 within the BPPE with a professional staff of three (UNDP, 1984: p.2). The CEO was tasked to:

- Develop and monitor evaluation policies and procedures. The new monitoring and evaluation (M&E) system was subsequently incorporated in the Programme Policy Manual for COs;
- Assist in transmitting results of project and thematic evaluations to all levels of staff;
- Collaborate with executing agencies and the Joint Inspection Unit in developing a consistent, coherent and agreed set of evaluation practices;
- Report on programme effectiveness by periodic assessment of project and thematic evaluations; and
- Assist operational units in developing countries to enhance their evaluation capacity.
A number of thematic and policy evaluations were conducted by CEO during this period while project evaluations were the responsibility of COs. In February 1994, as part of the ‘initiatives for change’, the Administrator established a new unit incorporating the CEO, named the Office of Evaluation and Strategic Planning (OESP). This combined evaluation and strategic planning unit would report directly to the Administrator, noting that:

The arrangement will strengthen the evaluation function in UNDP, a function to which I attach the highest priority, particularly with regard to follow-up on CEO recommendations. In addition to its evaluation functions, the new office will initiate strategic guidance and option analysis and development as they relate to UNDP themes and programmes. (UNDP 1994b: p.22).

The merger of evaluation and strategic planning functions was expected to boost the role of evaluation. But in reality, because of pressing strategic planning needs in the early years of UNDP’s reform, evaluation actually received less attention. Only a few corporate evaluations took place between 1994 and 1996. The expected strong feedback from evaluations to planning processes lacked corporate attention and a suitable mechanism to have meaningful influence within the organization.

But in 1997, OESP developed and introduced the new Monitoring and Evaluation Handbook, later known as the ‘Yellow Handbook’. This was the first detailed manual to provide step-by-step guidance for M&E at project and programme levels for use by all units. The book was considered part of the Programme Policy Manual. As a landmark output and a useful guide, it was well recognised by member countries and all stakeholders both within and beyond the organisation.

2.2 The dawn of RBM

Following a “major transformation in the very core constituents of the organization” in the 1990s,4 UNDP launched a reform project in 1997 known as ‘UNDP 2001’ (UNDP, 1999: p.4). As part of the initiative, the Administrator announced efforts to ensure the independence and effectiveness of the evaluation function (UNDP, 1997a). This would be achieved, in part, by establishing EO as a separate unit whose director would report directly to the Administrator on the results and performance of UNDP programmes.5 The strategic planning function of OESP would be placed in the Bureau of Management. The establishment of EO was welcomed by the Executive Board together with the clearer link between evaluation, lessons learned and programme enhancement (UNDP, 1997b).

In early 1998, a decision was taken by the UNDP Administrator to develop a framework for the measurement and assessment of programme results. It recognized that in the new environment of UNDP as a focused development agency rather than one transferring funds, it needed to stress results attained rather than the volume of assistance provided. UNDP also saw the emerging RBM paradigm as an opportunity to deepen an internal process of management reform begun in 1997 with the ‘UNDP 2001’ change initiative. The EO was assigned the lead responsibility to work on the design of such a framework. This step initiated the introduction of RBM in UNDP. As the discussion on future funding with the Executive Board matured into the first Multi-Year Funding Framework (MYFF) in 2004, EO became part of a unified management effort coordinated by the Bureau of Management to oversee RBM’s application.

---

4 Covering the mandate, vision, profile, funding base, resource allocation principles, mode of programme implementation and its relationship with the United Nations.

5 ‘UNDP 2001’ (DP/1997/16/Add.7) proposed that EO be housed in an office within the Bureau for Policy and Programme Support and that the Director report to the Assistant Administrator, but the Executive Board insisted that it be an independent office reporting directly to the Administrator.
In 2000, responsibility for coordinating the RBM process in the organization was assigned to the Operations Support Group (UNDP, 2000a), but successful implementation of RBM would also hinge on strong mechanisms for feedback and evaluation. It was believed that without a system in place to provide feedback on results, benchmarks could be produced and learning cannot take place. In this context, evaluation reports would provide more credible evidence through which UNDP substantive performance can be assessed.

The introduction of the RBM system in UNDP led to two key changes. First, evaluation was considered an integral part of the RBM system that emphasized results and substantive accountability. The concept of the ‘outcome’ became the cornerstone of project and programme design, and outcomes situated a particular project intervention within the wider country development context. The emphasis on effectiveness of contribution to an outcome in turn brought increased emphasis on partnership, synchronization and coordination in the process of design and implementation of initiatives.

Second, this shifted the focus of evaluation on project and programme results away from process, inputs and outputs. For the first time, evaluations emphasized the contribution of UNDP initiatives to in-country development results. The concept of outcome evaluation was introduced as a unit of evaluation within country programmes. The Yellow Handbook was substantially revised and revamped in 2002 to bring it more in line with the RBM approach. The essential emphasis on results analysis in CPEs emanated from this period.

A friend through the years: the M&E ‘Yellow Handbook’ produced by EO in its 2002 and 2009 iterations.

By this time, UNDP had an established M&E regime with a palpable organizational grounding. Although monitoring was not strictly an EO function, providing the guidelines was left with the Yellow Handbook for logical consistency and cohesiveness of approach. For this reason, the subsequent updating of the handbook to align it with the evaluation policy was done in 2009 jointly with Operations Support Group as the custodian of the monitoring function.

Predecessor evaluation units in UNDP assumed different institutional mandates and levels of independence in line with the organization’s evolving evaluation architecture.
2.3 The development of the UNDP evaluation policy

In 2004, EO volunteered to be the first subject of a peer review designed under the auspices of the Network on Development Evaluation of the OECD Development Assistance Committee (DAC). The review aimed to assess and enhance multilateral agency evaluation capacity and performance with an ultimate view of improving development performance. The subsequent peer review report answered the central question of “whether a multilateral agency’s evaluation system produces evaluations which are credible, valid and useable for learning and accountability purposes” (Cole, et al. 2005: p.3).

The peer review coincided with, and contributed to, the formulation of an evaluation policy for UNDP. EO management took this exercise as an opportunity to enshrine evaluation with an appropriate legislative mandate and situate evaluation within an organizational policy setting. EO carried out a review of evaluation policies of multilateral and bilateral organizations as well as regional evaluation associations. An extensive consultative process informed the new policy. Stakeholder meetings were conducted at the headquarters, regional, and country levels to assess perceptions of the evaluation function and to present principles of the emerging policy. The evaluation policy presented to the Executive Board in June 2006 had the following purpose (UNDP, 2006a: p.2):

“To establish a common institutional basis for the UNDP evaluation function. The policy seeks to increase transparency, coherence and efficiency in generating and using evaluative knowledge for organizational learning and effective management for results, and to support accountability.

Aligned with the norms and standards issued by the UN Evaluation Group6 (UNEG) in April 2005, the policy was intended to situate itself firmly within the priorities of Triennial Comprehensive Policy Review of Operational Activities for Development of the UN System (General Assembly resolution 59/250), which emphasized national leadership of evaluations and UN reform. It provided an opportunity to codify the various definitions and approaches to evaluation in UNDP. Most importantly, the policy established the institutional responsibilities among stakeholders for the evaluation function in the organization as a whole.

**Guiding Principles of the UNDP Evaluation Policy**

(a) Human development and human rights. Evaluation is guided by the people-centred approach of UNDP to development, which enhances capabilities, choices and rights for all men and women.

(b) UN system coordination and global partnership. Evaluation draws on and contributes to collaboration within the UN system to improve effectiveness and reduce transaction costs for development cooperation.

(c) National ownership. Evaluation should be guided by national priorities and concerns, including being inclusive of diverse national interests and values, strengthening partnerships with governments and key stakeholders, building national capacity to implement, monitor and evaluate.

(d) Managing for results. Evaluation supports UNDP to manage for results by assessing the extent to which UNDP’s processes, products and services contribute effectively to development results affecting people’s lives. Through this, evaluation highlights the need for quality in the design of programmes so that results are clear, measurable and can be monitored and evaluated. Through the generation of evidence, evaluation enables more informed management and decision-making for strategic planning and programming.

The policy provided for greater independence of EO, noting that the evaluation function should be “structurally independent from the operational management and decision-making functions in the organization so that it is free from undue influence” (UNDP, 2006a: p.3). To ensure independence, the policy first set appropriate standards for UNDP management and evaluators: “Management must not impose restrictions on the scope, content, comments and recommendations of evaluation reports. To avoid conflict of interest, evaluators must not be directly involved in policy-setting, design, implementation or management of the subject of the evaluation either before, during or after the evaluation” (ibid.). Second, it put more authority in the hands of the EO Director who could set the agenda for evaluation, after

---

6 UNEG is a professional network that brings together the units responsible for evaluation in the UN system including the specialized agencies, funds, programmes and affiliated organisations. It currently has 43 members and 3 observers.
due consultation, and submit reports directly to appropriate levels of decision-making without the need for clearance. The EO would report to the Executive Board through the Administrator and the appointment or removal of the EO Director required the endorsement of the Executive Board. Moreover, the EO Director would not be allowed to remain in UNDP after leaving the position.

In its approval of the first UNDP evaluation policy in 2006, the Executive Board asked EO to commission an independent review of the policy in three years. That was completed in 2010 (Gariba et al., 2010). The review was quality-assured by an independent panel of advisers and presented to the Executive Board (UNDP, 2010). It confirmed the continued relevance of the policy while highlighting improvements, namely in management responses and the use of independent evaluations. It was noted, however, that progress on decentralized evaluations was limited, particularly in their coverage, quality and compliance, despite the availability of guidance and support.

A revised policy approved by the Executive Board in January 2011 (UNDP, 2011) sought to clarify roles and responsibilities and to improve evaluation practice. Greater importance was given to national ownership in the evaluation process, national evaluation capacity development, and the use of evaluation for organisational learning and accountability. Most importantly, a new compliance requirement for decentralised evaluations was introduced intended to: (i) give flexibility to programme units in determining how to carry out evaluations; and (ii) promote joint efforts with national partners in evaluation. A second policy review was completed in 2014 and a revision to the evaluation policy was submitted to the Executive Board in 2015 for consideration. The review found that, based on a strong engagement with UNDP management and the Executive Board, EO was in a strong position to further strengthen its strategic role in UNDP (van den Berg 2013: p.2).

The evaluation function underwent significant reforms in the last 20 years in tandem with the wider organizational changes in UNDP. Each phase of reform provided an opportunity for EO to reassess its strengths and weaknesses to reposition itself for greater value. The publication of the Yellow Book as a component of programme policy, RBM’s institutionalisation, the OECD/DAC peer review, and finally the adoption of the evaluation policy cumulatively provided the impetus for evaluation to be embedded into the organisation. The interest of the Executive Board to increasingly utilise evaluations in their deliberations further legitimised an agenda for a credible and reliable evaluation function in UNDP.

### III. Country programme evaluation before the ADR

Throughout this evolution of UNDP’s evaluation function, the organization developed CPE approaches and practices. The ADR follows a series of experimental attempts at using the country programme as the unit of analysis in evaluating development results. This section chronicles early forms of country programme evaluation, notably the Country Level Impact Assessment (CLIA).

#### 3.1 Early efforts at country level review

The concept of mid-term reviews (MTRs) has existed as part of the country programming process since 1971. An evaluation of the MTR in 1990 found that the process had improved the quality of programmes but been focused on programme management, and suggested that MTRs focus on enhancing country programme effectiveness (UNDP, 1990: p.2). By the end of the 1990s, MTRs were required for all programmes over $10 million. The precise timing of the cycle, with a full assessment being conducted at the end of the fourth year to feed into the next programme. In reality, most of the MTRs were undertaken towards the fourth year.

---

7 As part of Decision 2006/20 the EB: “Requests that, as of the annual session 2009, UNDP submit a triennial review of the evaluation policy”.

8 In theory, the MTR would be conducted at the mid-point of the programme, towards the end of the second year of a five-year
review process and the involvement of other parties would be decided upon by the government in consultation with UNDP (DP/1998/19/Add.3). Many of these reviews were comprehensive and over time took an evaluative approach in their analysis with recommendations to be considered in formulating the next country programme. UNDP executive boards started taking active interest in country reviews in exercising their governance oversight of the country programmes.

With the corporate shift to RBM from the late 1990s, it was clear that country reviews had to become more results-oriented with increased emphasis on development results and partnership strategies. The assessment and validation provided by country reviews of UNDP support would thereby be central to the formulation process of new country programmes. Recognizing the need for reform, the custodianship for the reviews was transferred from the Operations Support Group to EO in November 2000 (DP/01/26). As the RBM system was being formally introduced, there was a parallel initiative on decreasing process load and requirements on COs. As a result, country reviews as a mandatory exercise were suddenly abolished. COs could still conduct reviews with RBx and EO cooperation where considered useful, but without submission to the Executive Board.

Country reviews served as a key management, accountability and learning tool for UNDP, governments, and by extension, the Executive Board, vis-à-vis UNDP engagement at the country level (DP/01/26). They were viewed as a way to report both on projects and the entire UNDP ‘presence’, such as its coordination role, advocacy, policy dialogue, and political initiatives. Although their scope was broad, in practice country reviews were unable to assess at the outcome level (Cossee 2001: p.4):

It is quite a light exercise: UNDP uses one or two independent consultants or (quite often) ex-staff members for about one month each to conduct a review. There are no beneficiary interviews and little fieldwork in general. The quality of the reports varies. Reliance on ex-staff members has been mentioned as a cause for the low quality of some reports.

3.2 Experimenting with country evaluation

UNDP experimented with independent CPEs beginning in the mid-1980s for Niger, Mali, and North Yemen (Reddy 1989, p.4). The 1987 Administrator’s annual report on evaluation noted:

Three independent evaluations of country programmes in three least developed countries (LDCs) have provided useful lessons for follow-up at the programme level and have provided experience in the conduct of these complex exercises. The lessons learned have been used to strengthen monitoring and evaluation in the country programme context. The sample to date is limited both in terms of programme size and complexity, and needs to be augmented by more evaluations (UNDP 1987).

CPEs were included in the Programme Policy Manual in early 1990s as an assessment tool for COs. Guidelines for country programme evaluations were used as the primary resource for experimenting new approaches with evaluations conducted over 1994-1996 in Sri Lanka, Niger, and Uganda (UNDP, 1993). These efforts to field test the guidelines were expected to be developed into a standard evaluation procedure for country programmes for circulation by the end of 1994 (UNDP, 1994a). These evaluations used the ‘3-M’ (macro-meso-micro) framework for their impact analysis, defined as follows:

- Macro impacts include upstream policy advice and development planning assistance that result in change of governmental policy and practice.
- Meso impacts are the intermediate translation of macro impacts into operational activities, improving government’s delivery capacities to the country.
- Micro impacts are the downstream direct provisions of goods and services to beneficiaries or the support and strengthening of self-help organizations.

---

9 In collaboration with the Swiss Government.
This phase of experimentation was short-lived, however (Cossee, 2001). The constraints for conducting CPE were multifaceted. As first generation CPEs, the OESP led these evaluations and as a result they were not demand-driven from the country level. To get a Resident Representative to volunteer for the CPE was difficult. UN staff were not receptive to being under the microscope, and some felt the process was a duplication of the MTR. There was also the difficulty of accepting the evaluation findings by resident representatives and sometimes by the RBx. Without policy legitimacy such as a mandatory clause, it would be difficult to mainstream CPEs in UNDP. Without any evaluation culture in place and an open mind to learn lessons from past experience, CPEs would remain a ‘lone ranger’ with occasional buy-in within the organization.

3.3 Piloting the Country Level Impact Assessment

In 1999, EO launched the first of three pilot CLIAs in UNDP’s initial attempt to assess long-term impact of its interventions in a country setting.10 CLIAs were conducted between 1999 and 2001 in Burkina Faso, Malawi and the Philippines aiming to:

- Document where projects, programmes and soft interventions supported by UNDP and other development partners had a significant impact;
- Identify the combination of external factors, capacities and management decisions that led to this significant impact to draw lessons applicable in a broader context and to advise UNDP on ways to improve its impact;
- Promote a result-orientation by focusing the attention of UNDP on the impact of its interventions and by building in-house mechanisms and capacities to evaluate impact; and
- Field-test the CLIA methodology, assess its efficacy and effectiveness in capturing outcomes and impacts, and recommend improvements.

UNDP defined impact evaluation within the context of CLIA as the broad, longer-term impact of a programme or outcomes, whether intended or unintended. Impact referred to the ultimate effect of a development intervention on people’s lives.11 The CLIA methodology’s application in the three pilots used an ‘interpretivist’ approach to measure impact through stakeholder groups. CLIAs evaluated stakeholder perceptions of UNDP interventions and whether, how and where they made a difference. This approach engaged participants in a dialogue to help convey an accurate portrayal of how people construct the meaning of their conditions. The use of both qualitative and quantitative inquiry methods and triangulation was especially important for the CLIA given the long-term nature of the information required to report on the development impacts. As a pilot study, CLIA experimented with the application of the various concepts and techniques used to identify and capture impact, while also aiming to understand their optimal management (e.g. planning, resourcing, scoping).

Results from the pilots suggested potential to assess UNDP’s development impact over a longer period. An EO workshop that reviewed the CLIA methodology and process found that generating and validating impact perceptions was possible using goal-free triangulation and participatory methodologies, and it provided an excellent way to capture unintended effects (UNDP, 2002b: p.26-27). UNDP interventions constituted primarily “soft” assistance (e.g. capacity building, advocacy, coordination, partnership building and resource mobilization) difficult to isolate and translate into micro-level tangible development impacts. This also suggested that CLIA could be inte-

---


11 It assumed that such a change can occur anywhere between 10 to 50 years after an intervention was first undertaken, although in practice CLIA looked at a time horizon of 15-20 years.
grated into a continuous results chain tracking system that links outcomes to impacts rather than be treated as a stand-alone exercise. The workshop highlighted that attributing impacts to UNDP was not feasible – especially given the wide scope, large volume, and long timeframes of its interventions – which suggested that CLIA could be better used as a joint country-driven impact assessment tool.

CLIA were, however, not adopted across the organization and remained an EO exercise without wider dissemination or recognition as a mainstream evaluation method. First, this was based in part on a deficit of interest of RBx and COs in historical or longitudinal impact assessment. Resident representatives were often more focussed on their current programme or recent past. Second, the results of the three pilots raised issues of the high transaction costs involved in applying the CLIA methodology. Among the lessons identified in the workshop, CLIA methods were found to be inherently costly and time consuming in UNDP’s typical context of high staff turnover and low institutional memory. The CLIA pilot exercise nonetheless revealed important lessons for UNDP on how to undertake country level assessment of performance. As will be seen in the next section, CLIA created a methodological awareness that shaped ADRs. On the other hand, for reasons described above, the CLIA equally generated a sense of reluctance in EO to engage in impact evaluation.

**IV. Rolling out the first generation (2002-2007)**

With the discontinuation of MTRs and country reviews as mandatory in 2000, the UNDP Administrator and the Executive Board had no way of tracking achievement or programme implementation after initial approval. While the country programme is the optimal unit for assessing UNDP’s contribution in a given country, there was no corresponding instrument of assessment. The EO was entrusted with the responsibility to develop an instrument that would support the Administrator’s substantive accountability while assisting future programme quality enhancement and generating organizational lessons for use.

EO started to work based on its experience in country evaluations to respond to this requirement.

The alternative system of CPE was initially called the Assessment of Country Results and later Assessment of Development Results. The UNDP Executive Team thus approved the concept of ADRs on 21 November 2001, and the first generation of ADRs were published and shared with the Executive Board in 2002. This section analyses the initial ADR approach and early changes in a period of crystallization.

**4.1 Original design and guidance**

Five ADRs were conducted in 2001 and 2002. They were based largely on the experience and exploratory initiative of EO management and senior staff. Programmes were assessed using a ‘goal-free’ approach, which provided room for creativity and innovation. Initial ADR guidance set out its purposes to (UNDP, 2002a):

- Support the Administrator’s substantive accountability function to the Executive Board and serve as a vehicle for quality assurance of UNDP interventions at the country level;
- Generate lessons to inform programming at the country and corporate levels; and
- Provide stakeholders in the programme country an objective assessment of results, specifically outcomes, achieved through UNDP support and partnerships with other key actors for a given multi-year period.

These three objectives may be reflected in any given country ADR depending on circumstances, as an opportunity for a CO to cement its position and vision vis-à-vis partners, as a tool for advocacy, learning and buy-in with stakeholders. The guidance also set out the ADR’s principles listed below.

---

Guiding Principles of the ADR in 2002

The ADR will:
- Be led by EO to ensure the integrity and independence of the evaluation.
- Provide a forward-looking analysis through an assessment of UNDP’s strategic positioning and development results.
- Focus on outcomes (i.e. changes in specific development conditions) and UNDP’s contribution to improve an understanding of the outcome’s status and influencing factors. The ADR does not assess impact.
- Look at the most strategic outputs delivered by UNDP to establish a credible link between UNDP’s output and the observed outcome.
- Be linked to corporate exercises and RBM by using accountability frameworks and M&E information as inputs.
- Be participatory by systematically documenting perspectives of all relevant stakeholders (e.g. CO, RBx, governments, donor community, NGOs and beneficiaries) to form a complete picture of UNDP’s activities and results.
- Analyse CO management issues on programme implementation only when such issues greatly influence the attainment of development results.

Source: UNDP 2002a: p.3

4.2 Scope and methodology

The initial ADR approach focused on key results, specifically outcomes, and covered all UNDP assistance irrespective of source of funding. Most importantly, because of its focus on outcomes, the assessment did not go into great detail on the contribution by specific projects. However, the ADR brought out important contributions of UNDP funds and programmes where relevant, and how UNDP works with these entities. When analysing outcomes, the evaluation considered both anticipated and unanticipated results, and positive and negative progress.

The ADR covered a given multi-year period, typically the last five years of programming including an analysis of intended results in future years. Consequently, the ADR straddled two country programmes and in so doing would provide an assessment of both UNDP’s strategic positioning and its contribution to development results. Analysis in the two areas is iterative, as the assessment of positioning will influence the appreciation of results, and vice versa. A combination of strong strategic positioning and performance on results signals a high value of UNDP to the country’s development efforts.

First, in terms of strategic positioning, the ADR focused on the added value that UNDP contributes in relation to those of its partners in addressing national development needs. This included a review of three key aspects, notably the UNDP country programme’s relevance to national needs and priorities; anticipation and responsiveness to changes in development context; and synergies and alignment with other partners. In sum, the ADR aimed to provide a succinct vision of how UNDP, given its comparative advantages, positioned itself in response to the surrounding environment, development priorities, and needs of stakeholders in order to recommend how UNDP could best position itself to leverage its unique strengths.

Second, the ADR ascertained UNDP’s contribution to significant development results in the country. This included both a descriptive element and assessment of how UNDP’s role contributed to progress towards intended results. The analysis of results pivoted around the areas of intervention of UNDP. Specifically, the ADR would:
- Identify major changes at the outcome level in national development conditions within UNDP thematic areas, and the factors that have influenced these results;
- Assess UNDP’s contribution to key development results within the thematic areas, including an estimation of the contribution of key outputs to the achievement of outcomes by reviewing UNDP programmes, initiatives, Country Cooperation Framework (CCF), and Strategic Results Framework (SRF);
- Assess anticipated progress in achieving intended outcomes in UNDP thematic areas, including based on the current status of intended outcomes; and

---

13 For example, UNDP may demonstrate good results in a specific area, but the area may however not be the most essential for future support. Alternatively, UNDP may be providing support in a very important area, but not be able to make a real difference with its contribution.

14 Now the Country Programme Document (CPD).
15 Now the Results Oriented Annual Report (ROAR).
Review UNDP development partnerships and how these contributed to results based on the nature of interaction with partners, level of influence to collectively generate outcomes, non-project initiatives to generate better partnerships, and comparative advantage to provide a neutral forum for partners to venture in soft and advocacy areas of reform.

Evidence was gathered through three major sources of information: perception, validation and documentation according to a concept of triangulation illustrated below.

**Triangulation methods**

**Perception**

Interviews with stakeholders

**Validation**

Statistical analysis of national data and indicators; field visits, direct observation; qualitative assessment of trends using secondary data sources

**Documentation**

Basic country documentation; donor analytical reports and programmes; M&E data, progress reports

*Triangulation methods as set out in the 2002 ADR guidance.*

ADR guidance set out four basic elements of the methodology. First, there was an emphasis on preparation and standard documentation. Various design strategies were applied depending on the thematic area and the nature of the outcome in order to answer different evaluation questions, or to dig deeper to answer one question better or more credibly. This preparation includes:

- Exploratory mission by the EO Task Manager to arrange local research and finalize the TOR with the CO management;
- Desk review with a programme mapping of UNDP focus areas and the linkages of key results frameworks (e.g. UNDAF);
- Local research through partnerships with institutions and/or consultants to perform a number of services such as: (i) logistical support; (ii) analysis of a thematic area; and/or (iii) data collection including surveys, beneficiary assessments, interviews, M&E reports, and baseline studies; and
- A mix of evaluative methodologies, including focus group interviews, sample surveys, questionnaires, mini-surveys, desk review/analysis of existing data, semi-structured interviews, statistical analysis and field visits.

Second, there were generally two broad approaches when assessing results at the country programme level, depending mostly on programme size. A top-down ‘subtractive’ process analysed overall achievements in the country, within a sector or thematic area, and then attempted to explain which parts of the national successes and failures are linked to the efforts of a particular donor. The ADR tried to attribute results to UNDP, not to specific UNDP-supported projects, although that would be an unavoidable corollary. A bottom-up ‘additive’ process took individual projects and aggregated findings using conventional evaluation techniques. This approach is time-consuming and most applicable when the country programme is relatively small.

Third, evaluators were to analyse programmes selectively through a purposive sampling, involving: (i) a large programme that in itself could be expected to have large repercussions influencing a strategic outcome (e.g. significant national poverty support programme); (ii) several projects linked to a specific outcome; or (iii) a project considered a pilot or innovative. These projects would be studied more closely, including analysis of project documentation, M&E reports, specific interviews and field visits. In addition to in-depth sampled projects, other projects may be visited according to selection criteria as outlined below.

### Criteria for project selection for field visits

A project/programme may be visited if:

- Expected to have large repercussions that influence a strategic outcome
- The only project linked to a specific outcome
- Considered a pilot/innovative project
- Clustered geographically with other projects with different outcomes
- Longstanding, such as four years or more
- Part of a very large programme
- Typified as a success story or perceived failure
- Reflective of the mapping of priorities and results
- Randomly selected for sample validation
Fourth, ADRs were to combine a ‘goal-free’ review of the previous programme cycle with a forward-looking assessment of the ongoing strategic framework. This meant that the ADR evaluated key UNDP past results, both achievements and failures, but did not aim to assess whether UNDP achieved what it planned. Once the status of achievements was determined, an assessment of UNDP’s progress towards pre-defined goals (as defined in the SRF) to answer if UNDP is on the right course to where it aims to be.

The ADR scope and methodology in its formative years was clearly rooted in the experience of country reviews, MTRs, and CIAs. RBM principles and practice helped to transform the approach to a more practical and applicable methodology combining the iterative nature of enquiry with issues of strategic positioning and development results.

4.3 Review and reform

The initial round of ADRs was later described as a pilot phase (UNDP, 2006b) through the adoption of a new evaluation policy in 2006 that changed the framework within which ADRs and all evaluations operate. The 2005 peer review of the UNDP evaluation policy (Cole et al., 2005) examined the UNDP evaluation function and found that the accountability and performance assessment purposes of ADRs were underexploited relative to those of learning. It therefore strongly suggested that the EO “re-examine and adjust the current heavy emphasis on learning from evaluation, to give greater weight to accountability” (Ibid: p.44). The ADR’s usage of goal-free evaluation was highlighted as limiting their utility for both accountability and learning. The review also found that ADRs were useful for strategic direction and informing debate by examining alternative approaches and lessons learned, consistently identified by operational managers as their major value. Finally, the review concluded that following initial experimentation in CPE approaches with the ADR, “a definitive model should be adopted as soon as possible” (Ibid, p.46).

The 2006 revised evaluation policy stipulated that ADRs should be conducted at the end of each programme cycle and made available to the Executive Board. For the first time, ADRs became a formal accountability tool (UNDP, 2006b). The policy made it clear that ADRs would be completed before the development of new country programme documents (CPDs) in recognition of a need to improve their timely input into programme design and decision-making. Objectives were added to ADRs, including: accountability to programme countries and the Executive Board; promotion of a culture of managing for results; knowledge and learning; and promotion of ownership and an M&E culture among national partners.

Together, these elements of review and reform contributed to enhancements and adjustments for the ADR in terms of formalizing its policy mandate, honing rigorous methodological guidance, drawing focus on substantive accountability, and eliminating the goal-free approach to assessment. UNDP’s overall evaluation architecture played a role in the evolution of the ADR. While the evaluation function’s importance varied over time, the 2005 peer review and subsequent 2006 policy mainstreamed ADRs as a mandatory exercise. These changes provided the foundation for a next phase to further scale production.

V. Acting on a policy mandate (2008-2015)

ADR entered into a second phase as a mandatory exercise at the onset of the 2008-2011 UNDP Strategic Plan, which laid out the role of evaluation in UNDP in line with the revised evaluation policy. This period saw an increase in the use of ADRs in thematic evaluations and meta-synthesis. While experimentation continued at a slower pace, there was still a series of discrete adjustments and additions that aimed to increase evaluation quality and use.
5.1 Scaling report volume

The 2002 ADR Framework and Guidance suggested that EO would conduct between five and 10 ADRs per year, and UNDP senior management endorsed six in 2002 and 10 in 2003. However, only four ADRs were produced per year on average until 2007. The 2005 peer review noted the potential to expand ADRs to cover all country programmes by producing two report types: (i) eight to 10 reports per year could use the standard ADR approach; and (ii) remaining country programmes could adopt a ‘light touch’ evaluation requiring fewer resources. The review cautioned, though, that an ambitious expansion of coverage would carry “intrinsic difficulties in maintaining and enhancing the quality of evaluations” (Cole et al., 2005: p.41):

As evaluators, we would counsel against diluting rigour and quality, especially of country programme evaluations, in order to increase coverage. This is a particular issue given the comparative lack of robust evidence available within UNDP’s systems upon which the EO’s evaluations can build. The Assessments of Development Results in particular are intended to be major evaluations of country programmes, and we would suggest that the aim should be to produce an appropriate number of well-selected and solid ADRs, with country offices and bureaux conducting other types of reviews to meet their management requirements across the remaining country programmes.

By 2007, there was pressure both internally from UNDP senior management and externally from some Executive Board members to increase the coverage to 100 percent of countries with programmes going to the Executive Board. The EO costed the scenario, but decided that its resources were too small to assume such coverage. Nonetheless, following the approval of the new evaluation policy, an increase in ADRs was planned for 2008. The figure below illustrates the rapid increase based on the year of publication.16

The institutional context of UNDP programming and other factors external to EO’s control have presented distinct challenges in the timely production of ADRs. When CPD periods are extended based on issues at the country level, an ADR may be presented to the Executive Board two or more years after it is conducted if not altogether cancelled. This has been one reason for a distortion in proportional ADR coverage of regions; cancellation of ADRs due to extended CPDs in the Africa region meant less coverage in a region with a large need for coverage. Evaluations have also been disrupted by conflict and instability (e.g. Egypt, Libya and Tunisia), in which cases ADRs are cancelled after preparations have already begun. This has presented a management issue in optimizing the use of human and financial resources.

The significant increase in ADRs conducted in 2008 was supported by an increased number of EO human and financial resources with the addition of three professional and two support staff. At the same time, EO changed its internal structure so that managers of thematic evaluations could also engage in ADRs, allowing more flexibility. Different approaches were also tried to increase to coverage within the staff constraints, including hiring companies to undertake the evaluations with mixed results.

---

16 Given the need for ADRs to feed into the country-level UN strategic planning process (i.e. CPD and UNDAF development), the ADR is typically conducted the year before it is presented to the Executive Board and published. The slight decrease in 2014 (Figure 2) is attributable to other programmatic evaluations – five regional programmes and the global programme – conducted by EO in the same year.
5.2 Consolidation of methodology

The increase in ADRs reiterated the need for clearer guidance to carry out the evaluations in a more consistent and rigorous manner for quality results. Until 2008, existing ADR guidance focused on process, not methodology. This was partly due to the nature of how ADRs were conducted. The process relied on external evaluation consultants to undertake the design based on principles and parameters set in the guidelines. However, there were inconsistencies in how ADRs were approached and even varying understandings of evaluation criteria.

In response, an updated ADR Manual in 2010 codified existing methodology to guide consultants in preparing ADR reports using a more consistent methodology, including clearer definitions of evaluation criteria and options for addressing different contexts. It provided guidance to EO staff and members of the ADR evaluation teams on the ADR substantive writing requirements (e.g. contents, criteria, methods and data), aiming to:

- Set standards for report writing and provide a basis for quality assurance and enhancement;
- Enhance rigour and credibility vis-à-vis COs, RBx, Governments and the Executive Board;
- Facilitate learning and synthesis of findings across ADRs;
- Disseminate good EO/international practices; and
- Provide a methodological reference for UNDP self-evaluation exercises at the country level.

5.3 Emphasizing national ownership

UNDP has a longstanding commitment to national ownership in evaluation. At one stage, evaluations could only be undertaken at the behest of the programme country government (UNDP 1970: p.42): Evaluation carried out within the United Nations system of UNDP-assisted activities will only be made with the agreement of the Government concerned. This will be carried out jointly by the Government, UNDP, the United Nations agency concerned and, where appropriate, the executing agent outside the United Nations system. Such evaluation will be on a selective basis and restricted to the minimum essential for the improvement or follow-up of the projects concerned for the needs of Governments and for the improvement of the Programme. With the agreement of the Government concerned the results will be communicated, to the Governing Council for its information.

As early as 2002, the Executive Board made a decision encouraging UNDP to “ensure the full participation of the national authorities in drafting the terms of reference, the development of evaluation methodologies and indicators, and the selection of evaluation teams in all evaluation exercises conducted at the country level” (UNDP, 2003). The importance of national ownership goes beyond simply compliance with a fundamental UN principle of engagement. Its purpose is to ensure high quality of the product and most importantly ensure the evaluation’s utilisation by the national partners, as noted in the 2005 peer review (Cole, et al. 2005: p.52):

Achieving the necessary degree of transparency and consultation with all stakeholders at all stages of evaluations is demanding, but without it the base for ownership may not have been built, the take-up of recommendations is likely to be weaker, and indeed the due process and good faith in the process can be called into question.

However, efforts to ensure national ownership in ADRs fell short. In reviewing UNDP’s evaluations, the same 2005 peer review found “relatively little emphasis to date on the involvement and ‘ownership’ of partner country stakeholders in these evaluations beyond serving as interviewees, and participants in follow-up workshops in the cases where these are held” (ibid, p.42). In response, the revised 2006 evaluation policy integrated national ownership as one of its four key principles, noting (UNDP 2006a: p.3):

Evaluation should be guided by national priorities and concerns. It should be inclusive and take into account diverse national interests and values. Evaluation should strengthen partnerships with governments and key stakeholders. It should build the capacity of national institutions to implement, monitor and evaluate.

The Executive Board decision approving the Evaluation Policy specifically requested UNDP “to conduct evaluations of its operations at the country level, in close consultation with national governments” (UNDP 2006c). In response, the Annual Report on Evaluation for 2006 noted one of the key priorities for the year was “to engage key stakeholders of the organization and national partners in implementing
the policy” (UNDP, 2007a: p.3). An EO workshop designed in the context of the Executive Board decisions in 2007 led to the establishment of national reference groups as mechanisms for providing inputs to the design, specifically the terms of reference, and to review early findings of the team (UNDP, 2007b). Increased engagement throughout the process was assumed to enhance utilization and strengthen the full understanding of context. Reference groups were not made compulsory. The idea is proposed to government counterparts during preparatory missions and can be accordingly tailored. At a minimum, membership consists of government officials, but most national reference groups also include civil society members and key donors. In some cases, the group is used as a sounding board for conclusions and areas of recommendation.

While ownership of the evaluation process and results is traditionally oriented towards governments, an inclusive evaluation process ensures that ADRs capture the views of a broad range of stakeholders. Government counterparts have typically a bigger stake in the evaluation process and retain the responsibility to implement any recommendations emerging from the evaluations. For this reason, ownership from their position as policy-makers is critical. In addition, non-government stakeholders, beneficiaries and disadvantaged groups can be more valuable in providing evidence on benefits and results of UNDP’s support. An inclusive process of evaluation encompasses and solicits perceptions on an intervention irrespective of their status in the power and influence orbit. ADR teams must therefore be able to permeate the sense of engagement for these respective groups in the country.

5.4 Changing nature of the consultants

The nature and role of external members in ADR teams have changed significantly over time. Early efforts even included UNDP staff members in the evaluation teams, including the 2002 Sudan ADR that used the Assistant Resident Representative from Haiti and a staff member from the Bureau for Crisis Prevention and Recovery. From the start, EO recruited and managed a team of independent external consultants, led by an external Senior Team Leader. EO’s role was central in launching the exercise, formulating the TOR, managing internal consultations, and ensuring report quality and standards. EO exercised this role through the Task Manager who would be fully responsible for managing the process and the consultants. In some cases, the EO Task Manager acted as a full member of the team by participating in data collection, analysis, and report writing. In others, the EO manager provided little substantive input beyond comments raised during the report’s review. Although no formal assessment was made on the use of external consultants as team leaders, the feedback of products was generally of high professional quality.

As early as 2007, EO contracted international consulting companies to undertake ADRs to increase efficiency and address the shortage of staff. The experience was mixed in terms of efficiency gains and proved in some cases to be more expensive. The approach was not repeated and no official assessment was made of its value. In the move towards greater use of national consultants, there were however some examples of national institutions being contracted to undertake the evaluations.

In 2011, EO began considering a model in which its staff would lead evaluations by participating in all aspects of the ADR, including design, data collection, analysis and report writing. In theory, this would facilitate further cost reductions and promote consistently high quality products. It also reflected the changing nature of the EO staff profiles with more members being experienced evaluators in their own

---

17 ADR Benin and ADR Republic of the Congo.
right rather than the mangers of evaluation consultants. Concerns of EO task managers about a projected high work burden resulted initially in the approach being made optional, and consequently there was limited adoption in the 2012 round of ADRs. By 2013, the EO Director decided to make the model mandatory for all future and ongoing evaluations. The title Task Manager was changed to Evaluation Manager to indicate the increased responsibility.

5.5 Enhancing corporate alignment and use

ADRs have more recently introduced several changes to improve harmonization with UNDP’s vision and corporate systems for stronger evaluation use and rigor. First, it has provided a higher priority to management responses in the evaluation reports. A management response system has been in place since the start of the ADR and the Annual Report on Evaluation tracks compliance with respect to their preparation and implementation. However, these were not included or highlighted in the ADRs. In 2014, ADRs began annexing the full management response and including a summarized version in the executive summary with related recommendations. In parallel, IEO encouraged COs to present the full management response at the ADR stakeholder workshop in a similar way, which represents the report’s presentation in-country as shown in the timeline below.

**ADR annual average timeline**

In contrast to previous practice, the tabling of a draft management response with stakeholders (e.g. government, civil society and international development partners) would lead to more concrete and realistic actions. It was also believed that it could increase the likelihood of partner support of implementation of the recommendations for the design of the new CPD. The ADRs feed directly into the process of preparing the CPD as a change process where new ideas and directions can be directly adopted. To maximise use, actions in addition to the stakeholder workshop are undertaken, for example translation into the national language and in-country dissemination.

A second major area of corporate alignment has been the ADR revision in line with the 2014-2017 Strategic Plan. One of the principles of the first generation of ADRs included alignment with planning instruments, and efforts were made during earlier revisions of ADR guidance to facilitate the alignment process. In practice, alignment was lacking in a number of key areas. A 2013 meta-synthesis of ADRs found that “the metrics used by management to assess performance of the country programmes...are different from those used by IEO to evaluate performance under the ADRs” (UNDP, 2013b). IEO responded by aligning ADR methodology with key Strategic Plan programmatic and operational changes, including coverage of issues related to UNDP’s vision to reduce poverty, inequality and exclusion as well as parameters of how it manages quality programmes.

As IEO staff increasingly led ADRs and data collection, it became clear that there was a greater availability and quality of national data and UNDP monitoring data than when the ADR was initially designed. In many countries, there was also an increase in the coverage of decentralized evaluations and according to IEO’s quality assessment, an overall increase in quality too. As a result, ADRs moved further from the original emphasis on interviews towards greater use of other sources of evidence. These changes in UNDP’s corporate vision and systems have been accompanied by a larger reform agenda within the UN Development System. For ADRs to remain useful and relevant, they must continue to take into account

---

18 The Annual Report on Evaluation analyses the evaluation function in UNDP and its associated funds and programmes at the Executive Board annual session in June.

19 This practice was subsequently also adopted in IEO thematic evaluations.

20 IEO reflects the name change footnoted above.

21 IEO evaluations equally revealed a lack of alignment of country programmes with the Strategic Plan (UNDP 2012).
VI. ‘Evaluating as one’ in the UN system

This section explores the status of CPE in the UN development system and UNDP’s appropriate role given its experience with ADRs. It seeks to provide a forward-looking assessment of how UNDP can provide leadership to invigorate CPE in line with the direction of wider UN reform.

6.1 Country programme evaluation behind UN reform

Efforts to improve CPE of UN partners have historically lagged behind UN reform to forge a more coherent UN system at the country level. Over a decade after Secretary General Kofi Annan announced his reform plan in 1997, the 2008 Triennial Comprehensive Policy Review in 2008 provided impetus to “optimize the link of evaluation to performance in the achievement of development goals” and to strengthen the UN evaluation activities through the UNDAF (United Nations 2008). In 2010, new guidelines were introduced that made UNDAF evaluation a mandatory exercise in the penultimate year of all UNDAFs (DOCO 2010). In parallel, the Secretary-General launched Delivering as One (DaO) in an attempt to make the UN system more coherent, effective and efficient. Since eight countries volunteered to become DaO pilots, over 40 have formally adopted the approach and more are actively implementing its core aspects. The approval of the Standard Operating Procedures (SOPs) in 2014 outlining DaO country procedures has equally made evaluations of DaO countries mandatory.

In this context, both DaO and UNDAF evaluations, herein referred to as UN joint country evaluations (JCEs), are no longer a peripheral component of UN reform. At the global level, there is growing political momentum to push forward a more integrated UN country development system. All UNDG members have approved the SOPs, and each UN Executive Head has committed to promoting implementation actively through follow-up and collective monitoring. UNDG Chair and UNDP Administrator Helen Clark stated that every UNCT should be “taking steps towards the progressive implementation of the SOPs in their country” (UNDG, 2014). Reform has remained a central agenda item in 2015 with the 70th Session of the UN General Assembly on the post-2015 agenda and ongoing efforts to transition from the 2012-2016 cycle of the Quadrennial Comprehensive Policy Review (QCPR). As outlined below, there are an array of UN partners in addition to UNDP carrying out CPEs, which nominally overlap in evaluating shared UNDAF strategic frameworks.

<table>
<thead>
<tr>
<th>Country programme evaluations conducted by UN partners, 2004-2015&lt;sup&gt;24&lt;/sup&gt;</th>
<th>UN partner</th>
<th>CPE</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAO</td>
<td>Evaluation of Cooperation in Country</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>GEF</td>
<td>Country Portfolio Evaluation</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>IFAD</td>
<td>Country Programme Evaluation</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>ILO</td>
<td>Country Programme Evaluation</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>UNDP</td>
<td>Assessment of Development Results</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>UNFPA&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Country Programme Evaluation</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>UNICEF</td>
<td>Country Programme Evaluation</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>UNIDO</td>
<td>Independent Country Evaluation</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>WFP</td>
<td>Country Portfolio Evaluation</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

<sup>24</sup> Among several purposes, these evaluations were framed as a way to guide the UNCT’s strategic planning of the subsequent UNDAF cycle.

<sup>25</sup> In 2015, over half of UNDP’s network of country offices in the Africa region were DaO.

Note that this presents CPEs of UN partners who participate in the UNDAF process, and does not include country reviews or assessments carried out by other UN entities. Data derived from UNEG, FAO, UNDP, IFAD, UNICEF, UNIDO, ILO, and UNFPA as of September 2015.

<sup>26</sup> UNFPA uses a ‘hybrid’ model: CPEs are commissioned by the central evaluation unit, but conducted in a decentralized manner, in contrast to other UN evaluation units.
However, there is a persistent disconnect between the expectations set for UN evaluation at the country level versus actual implementation. JCEs continue to be underfunded given their ambitious scope and focus more on UN coordination mechanisms as ends in themselves than on contribution to development results. UNDAF evaluations have an average budget of about $30,000\textsuperscript{26} compared to the recommended minimum of $100,000 for “conducting a high quality evaluation process in line with UNEG Norms, Standards and Ethnical Guidelines” (UNEG, 2011). Moreover, there has been weak compliance and poor report quality. A study of UNDAF evaluations commissioned by the UN Independent System-Wide Evaluation Mechanism notes that only 33 of 90 programme countries (about 37 percent) with active UNDAF cycles submitted evaluations in line with the 2010 UN requirement (Tristi, 2015).

The governance structure to facilitate greater quality and compliance in producing JCEs is very weak. There is no central repository, no sufficient means of quality assurance for their independence, and no follow-up mechanism to monitor implementation of recommendations. To date, DaO evaluations have been carried out in more than five DaO pilot countries, which largely mirror the approach to UNDAF evaluations while closely analysing harmonization and coherence factors that have affected the UNCT’s contribution.\textsuperscript{27}

In this context, UN coordination mechanisms in the field of evaluation have not been able to galvanize UN partners to provide coherent coordination of CPEs. While other many aspects of UN reform have advanced, changing the nature of evaluation at the country level remains an afterthought. UNEG has made genuine efforts to set up a Joint Evaluation Task Force in 2010 and publish related materials for conducting such exercises (UNEG, 2013), but its role has been relatively marginalized to updating methodological guidance. The Task Force has since been dissolved, and follow-up of use of its materials for joint evaluation has not happened. Amid the cautious approach to reform in the UN evaluation community and the political constraints of operating across multiple agency mandates, operational timelines, and visions, UNEG has not been able to provide the momentum to consolidate a clear vision of how CPE should be coordinated at the required quality standards.

Aligning the ADRs to these processes is a major challenge. In the face of the emerging UN reforms affecting country level programming and implementation, the continuation of ADRs as a systematic organizational exercise will need to carefully reassessed and balanced by UNDP.

\textbf{6.2 ADRs in an integrated UN development system}

On the one hand, UN agency-specific CPEs such as ADRs may face decreased demand and relevance given the increasing role of JCEs. As the UN has moved towards planning, programming and evaluating as one, UN agency CPEs confront a more crowded evaluation environment. There is greater possibility for overlap and duplication with JCEs as well as inefficient use of resources. This may reduce demand for products such as the ADR at the country level, although not necessarily from stakeholders at the headquarters level. Moreover, UN agencies must evaluate their contribution to shared outcomes and joint programmes within UNDAF frameworks. A more integrated programme should come with more integrated evaluation to stay relevant and credible.

\textsuperscript{26} Based on the list of final UNDAF evaluations from the Evaluation Resource Centre (erc.undp.org), accessed Jun 2015 2015.

\textsuperscript{27} See, for example, the country-led evaluations of the DaO programmes carried out in 2010 in Rwanda, Albania, Tanzania, Vietnam and Cape Verde.
On the other hand, the quality, scope and governance of JCEs limit their usability for traditional UN headquarters evaluation functions. First, JCEs consist of small-scale exercises in contrast to ADR-style products. They are not able to capture and rigorously assess UN agency development contributions at the country level. Even if JCE quality was improved, they have not provided a holistic picture of UN agency performance in line with current evaluation standards and functions at the headquarters level (e.g. accountability to the UNDP Executive Board and learning at corporate levels).

If JCEs are to displace ADRs, this suggests that either JCEs need to elevate their standards, or expectations of the level of rigour and detail provided in a traditional CPE will need to be recalibrated. To achieve the former, the UN system would need to increase resources and efforts to significantly improve JCEs. This would certainly require a significant shift to strengthen financing mechanisms currently in place, which generally rely on joint funding by evaluation commissioners, UNCT and national government. The incentives to do so will vary based on political demand of partners, and in this sense a stronger political mandate from the UN Development Operations Coordination Office (DOCO) will be a key factor.

UN partners carrying out ADR-like products are typically few, and the notion of standardizing evaluation for UN agencies with different mandates has created pushback at the headquarters and regional levels. In the past, IEO has offered to provide money from its ADR budget for UNDAF evaluations with no success. This reinforces the perception that while the UN system continues to integrate certain core functions, including programmes and planning, evaluation has lagged behind as an agency-specific function.

It is unlikely in the short-term that JCE quality will significantly improve, but it is probable in the medium to long-term that the ‘evaluating as one’ agenda will increase in importance to the point of displacing ADR-style products. In the near term, UNDP should accordingly continue conducting ADRs while refining internal guidance on how ADRs should work with JCEs and UN partners. This will help IEO lead and coordinate with partners for the good conduct of JCEs. ADR managers may, for instance, consider scenarios to: (i) share analytical work with the Resident Coordinator Office for inputs into the JCE; (ii) conversely, synchronize some data collection activities of evaluation teams; and (iii) conduct a joint evaluation with a fully integrated team.

### 6.3 The spirit of increased coverage

As ADRs are increasingly integrated into mandatory UN evaluation processes at the country level, there is an opportunity to increase evaluation coverage of country programmes in tandem. In fact, UNDP Executive Board members have consistently called for an increase in the proportion of countries covered by ADRs during each programme cycle. This recognizes that ADRs provide a valuable accountability and learning tool, but lack the sufficient coverage to provide a holistic picture of UNDP’s performance. Indeed, coverage has historically been on average about one-third of country programmes going to the Board. As shown in the table below, this figure varies significantly by year depending on the number of

<table>
<thead>
<tr>
<th>Year</th>
<th>Programmes going to the Board</th>
<th>Programmes covered by ADR</th>
<th>ADR coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>29</td>
<td>3</td>
<td>10%</td>
</tr>
<tr>
<td>2014</td>
<td>12</td>
<td>7</td>
<td>58%</td>
</tr>
<tr>
<td>2013</td>
<td>13</td>
<td>2</td>
<td>15%</td>
</tr>
<tr>
<td>2012</td>
<td>26</td>
<td>11</td>
<td>42%</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>23</td>
<td>29%</td>
</tr>
</tbody>
</table>

28 For example, the 2014 Tanzania ADR collaborated with IFAD and WFP who were conducting country-level evaluations in the same year.

29 Most recently during the June 2015 Executive Board session, the representative of the Government of Norway stated on behalf of other Executive Board members (Netherlands, Sweden, UK, and Canada): “It appears useful if IEO and UNDP management could develop some type of practical strategy of more comprehensive evaluation coverage of the country programmes at their conclusion and prior to the new country programme submissions to the Board.”
country programmes submitted, as well as by un-anticipated CPD extensions that can increase the volatility of coverage.

The practical implications of an increase in coverage would be far-reaching. It could provide the organization with a statistically significant, independent assessment of performance. This would represent an enhancement of UNDP’s performance assessment and overall RBM system. In its current form, UNDP assesses organizational-wide results in its Annual Report of the Administrator on the Strategic Plan in a ‘development report card’. At the outcome and impact level, performance data is derived from internationally published data sources that have a tenuous causal linkage with UNDP’s contribution. At the output level, performance data is derived and validated largely by UNDP self-reporting, namely the ROARs and decentralized evaluations, whose veracity and independence have been found to be seriously flawed (Baastel, 2014). As a result, there is a gap in the availability of impartial evidence to measure UNDP’s contribution to development results. In this context, an increased coverage of ADRs of programme countries would help to fill such a gap by requiring most or all country programmes to undergo a formal and independent assessment in addition to the current range of monitoring and reporting tools.

Nonetheless, achieving such an increased coverage at a ‘reasonable’ cost without compromising report quality and independence remains a challenge. As described in this paper, IEO peaked its yearly ADR output in 2009 and 2012 with 14 and 12 reports respectively. This represented a significant strain on staff and resources. In the current ADR model, which has shifted to IEO staff fully managing the ADR cycle since 2012, the production of over 10 ADRs per year represents an even greater burden on staff resources. A number of strategies to increase coverage are presented in Annex 2 and recognize these inherent challenges. These approaches are not mutually exclusive and are broken down in terms of projected trade-offs compared to the current ADR model, notably the estimated levels of quality, resource efficiency, and scalability to reach full coverage of UNDP programme countries.

When taking account of the fluctuation of the number of country programmes going to the Executive Board per year and, therefore, the number of programmes eligible for an ADR, a mix of these strategies could be envisioned. For example, IEO may elect to adopt a ‘light ADR’ approach in a year with 10 eligible countries, but scale-up production using both a cluster and hybrid management approach in a year with 25 or more eligible countries. This assumes that IEO would have sufficient lead time from one year to the next to prepare its strategy in advance without overstretching resources. It would also suggest greater budgetary variability based on the projected yearly coverage. Finally, other external factors may improve IEO’s ability to scale-up coverage. For example, at a stage when there is an increased focus on improving UNDP’s decentralized evaluation function, ADRs may in future benefit from stronger quality project, outcome, and country evaluations produced by commissioning offices.

---

30 Figures based on the online document list of Executive Board sessions: www.undp.org/content/undp/en/home/operations/executive_board/documents_for_sessions/
31 While IEO plans the amount of ADRs depending on the number of country programmes expected to be submitted to the Board, CPD periods are often extended and as a result delays when ADRs are presented to the Executive Board. For example, ADRs completed in 2012 may need to wait for a delayed CPD submission in 2014-2015.
32 The report card provides an overview of output performance by calculating average percent achievement according to progress towards set milestones. See DP/2015/11.
33 See a detailed methodological description in DP/2015/11/Annexes.
VII. Conclusions and lessons

This paper has documented UNDP’s evolving institutional context and evaluation function influencing the emergence of ADRs. In its infancy, UNDP’s evaluation architecture focused on project implementation, and even disbanded its independent division in 1973 in favour of merging with strategic planning and policy units. An independent unit was re-established in the late 1990s and evaluation exercises deepened assessment methods using the country programme as the unit of analysis. The initial cohort of ADRs in 2002 came at the introduction of the RBM system in UNDP in which evaluation assumed a more prominent institutional role and shifted focus on programme results away from a focus on process, inputs and outputs.

During this period, an extensive range of UNDP assessment, review and evaluation products at the country programme level have grappled with similar methodological challenges inherent in assessing aggregate-level results. MTRs and CLIs, like ADRs, faced the wide scope, large volume, and long timeframes of UNDP’s interventions. ADRs continue to undergo improvements in rationale, approaches and methodologies in search of rigorous evidence about UNDP’s development results in its programming countries.

In conclusion, this paper sees the future of ADRs and joint UN evaluation at the country level as closely tied together. The evaluation community can draw lessons from both ADRs and CPEs of other UN partners, in order to further strengthen JCEs. First, UN partners who are conducting CPEs can work through UNEG to strengthen JCE guidance. In contrast to the current JCE model of outsourcing one evaluation report, UNEG guidance may for example propose multiple approaches depending on the UNCT context. Second, UN partners should empower DOCO to strengthen JCE governance, including by setting up an appropriate funding mechanism. DOCO will be an important actor in consolidating support among UN partners and rolling out any mechanisms for ‘evaluating as one’ to ensure that UN country programmes conduct quality, timely and independent JCEs. In this respect, DOCO must also improve the governance structure of how JCEs are archived, quality assessed, and monitored. Such a system could be based on current platforms of UN partners in which evaluations are uploaded, selected for quality assurance, and have their management responses tracked publicly.

Several additional lessons may be gleaned specifically from ADR experience for application to the future integration of evaluation in the UN system:

1. **Find the right independent model** The model of IEO staff leading ADRs follows over a decade of experimentation with evaluation teams led by external consultants. For JCEs to be useful at the headquarters and country level, they will need to adopt contracting mechanisms that produce credible reports through independent teams while remaining linked with country-level RBM systems.

2. **Provide forward-looking analysis** The primary users of ADRs are the programme countries focusing on strategic inputs into the next CPD. Since the 2005 peer review suggested a greater emphasis on accountability, ADRs have continued to calibrate a mix of elements of accountability and learning with a focus on country-level users. Forward-looking analysis and recommendations remain integral to these evaluations, and a more systematic JCE product would equally provide formative value in programming for UNDAFs and DaO integrated frameworks.

3. **Focus on outcomes** The first generation of ADRs used a goal-free additive approach and were often more think pieces. ADRs became more oriented towards outcomes over time, which has presented a new set of challenges. Despite these issues, outcomes should continue to be at

---

34 For example, UN partners who are the ‘lead partners’ of certain UNDAF outcomes could also lead the outcome-focused evaluations of those respective areas to be synthesized into a larger evaluation report.

35 See, for example, UNDP’s Evaluation Resource Centre: erc.undp.org

36 For example, outcome frameworks may simply not be used by the CO with a weak linkage to the ensemble of projects actually being implemented.
the core of the JCE process and analysis given the structure of integrated UN planning.

4. **Enhance contribution analysis** Due to their focus on outcomes, ADRs evaluate programmes using broader outcome statements at a higher UNCT level, which has further compounded issues of causal attribution. ADRs use an outcome-based approach that takes into account contextual factors and other internal/external variables in assessing UNDP’s contribution within the UN system. JCEs must equally face a complex network of actors in their assessment of integrated frameworks. While attribution remains the gold standard, JCEs will need to refine their capacity to conduct contribution analysis beyond a focus on coherence mechanisms and outputs.

5. **Ensure participatory, inclusive evaluation** The ADR process has remained participatory and sought to formalize consultative mechanisms for both government and non-government stakeholders in order to promote ownership and ultimately use. JCEs should ensure that ownership constitutes its core set of principles, and furthermore that they provide a real means to include all segments of society in providing feedback on UN performance.

Notwithstanding the value of more integrated UN evaluation at the country level, there are in practice enormous challenges to moving forward. Each UN partner conducts CPEs within a distinct evaluation mandate, scope, and budget. While UNCT members share outcomes articulated in the UNDAF, the timing and approaches of CPEs are not fully harmonised. These challenges are evidenced by UNDP’s attempts to integrate its Moldova ADR with an UNDAF evaluation in 2011. The two exercises were conducted in a synchronised manner by one team of evaluators producing two distinct deliverables. Such efforts were not replicated partly due to issues related to the transaction costs associated with operational and methodological coordination.

For this reason, continued experimentation with JCE is needed, including collaboration between the ongoing CPEs of various UN partners and UNDAF or DaO evaluation efforts. New models need to address lessons learned from Moldova and other cases building on work of the Joint Evaluation Task Force. Both UNEG and DOCO have a greater role to play in providing the leadership to face potential agency resistance of joint evaluation of their work at the country level. Equally, the political mandate for ‘evaluating as one’ will be key to strengthening what will likely prove a difficult process of more closely aligning various CPEs. As the UN Development System continues to integrate its strategic planning and operational functions, evaluation should not continue to trail behind if it is to provide a more robust assessment of how the UN is affecting people’s lives.
Annex 1. Documents Consulted


### Annex 2. Strategies for increased ADR coverage

<table>
<thead>
<tr>
<th>Option</th>
<th>Approach and scalability</th>
<th>Resource efficiency</th>
<th>Quality / independence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Current ADR model)</td>
<td>IEO staff fully manage each ADR (e.g. preparatory, data collection and final missions). Use of ‘principal consultants’ allows some flexibility to manage multiple ADRs but still requires significant staff input and oversight. Approximately 5-10 ADR countries per year.</td>
<td>Current ADR unit cost is c. $160,000.</td>
<td>High/high IEO centralised management enhances quality / independence.</td>
</tr>
<tr>
<td>Hybridize management of process</td>
<td>A model similar to UNFPA in which ADRs are conducted by independent external evaluators who are ‘pre-qualified’ by a central evaluation unit but managed by the business unit being assessed. 20+ ADR countries per year.</td>
<td>High A large potential cost reduction (decrease in mgmt. costs, staff, and int’l travel).</td>
<td>Medium/Low Quality control over product is limited to TOR and/or other design elements. High risk of bias.</td>
</tr>
<tr>
<td>Cluster countries in a combined ADR</td>
<td>IEO could manage an ADR covering multiple country programmes with similar regional, programme or developmental traits. This approach has been used in the past by IEO in the Gulf States and is currently used by the World Bank Independent Evaluation Group. 15+ ADR countries per year.</td>
<td>Medium Cost savings may be significant with reduced travel, increased mgmt. efficiencies.</td>
<td>Medium/High Difficult to have context-specific analysis / recs but potential increased strategic value.</td>
</tr>
<tr>
<td>Reduce ADR scope and/or lighten methods</td>
<td>While current ADRs cover up to two programme cycles and the entirety of UNDP country activities, a “light ADR” could assess UNDP contributions through scaled-down scope, missions, limited project selection, revised methods, and more concise output (e.g. 30-40 pages instead of 80-100 pages). 15+ ADR countries per year.</td>
<td>Low Cost reductions may not be significant without major tweaks to current ADR model.</td>
<td>Medium/High Maintains IEO quality control, but risk of more superficial assessment of only certain programme components.</td>
</tr>
<tr>
<td>Increase joint evaluations</td>
<td>IEO places increased emphasis on participating in, and using, joint UN evaluations of shared UNDAF outcomes at the country level (e.g. DaO and UNDAF evaluations) as major inputs into ADRs. 10+ ADR countries per year.</td>
<td>Low High ‘coordination burden’ to ensure timing and good quality of inputs.</td>
<td>Low/Low Not currently feasible based on the dearth and low quality of these evaluations.</td>
</tr>
</tbody>
</table>

---

37 UNFPA Evaluation Policy, 2013  
38 In 2008, IEO conducted an ‘Evaluation of the Role of UNDP in the Net Contributor Countries of the Arab Region’, which began as a clustered ADR and then was hybridised additionally into a strategic evaluation.  
39 The IEG conducts Clustered Country Programme Evaluations (CCPES), including for example one of Resource Rich Developing Countries (Bolivia, Kazakhstan, Mongolia and Zambia) in 2014. See: ieg.worldbankgroup.org
## Annex 3. List of ADRs

<table>
<thead>
<tr>
<th>#</th>
<th>ADR Country</th>
<th>Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tanzania</td>
<td>2015</td>
</tr>
<tr>
<td>2</td>
<td>Armenia</td>
<td>2014</td>
</tr>
<tr>
<td>3</td>
<td>Malaysia</td>
<td>2014</td>
</tr>
<tr>
<td>4</td>
<td>Iraq</td>
<td>2014</td>
</tr>
<tr>
<td>5</td>
<td>Uruguay</td>
<td>2014</td>
</tr>
<tr>
<td>6</td>
<td>Algeria</td>
<td>2013</td>
</tr>
<tr>
<td>7</td>
<td>Sierra Leone</td>
<td>2013</td>
</tr>
<tr>
<td>8</td>
<td>Timor Leste</td>
<td>2013</td>
</tr>
<tr>
<td>9</td>
<td>Afghanistan</td>
<td>2013</td>
</tr>
<tr>
<td>10</td>
<td>Kenya</td>
<td>2013</td>
</tr>
<tr>
<td>11</td>
<td>Angola</td>
<td>2013</td>
</tr>
<tr>
<td>12</td>
<td>Croatia</td>
<td>2013</td>
</tr>
<tr>
<td>13</td>
<td>Cote d’Ivoire</td>
<td>2013</td>
</tr>
<tr>
<td>14</td>
<td>Niger</td>
<td>2013</td>
</tr>
<tr>
<td>15</td>
<td>Liberia</td>
<td>2012</td>
</tr>
<tr>
<td>16</td>
<td>Egypt</td>
<td>2012</td>
</tr>
<tr>
<td>17</td>
<td>Sri Lanka</td>
<td>2012</td>
</tr>
<tr>
<td>18</td>
<td>Paraguay</td>
<td>2012</td>
</tr>
<tr>
<td>19</td>
<td>United Arab Emirates</td>
<td>2012</td>
</tr>
<tr>
<td>20</td>
<td>Democratic Republic of Congo</td>
<td>2012</td>
</tr>
<tr>
<td>21</td>
<td>Djibouti</td>
<td>2012</td>
</tr>
<tr>
<td>22</td>
<td>Tunisia</td>
<td>2012</td>
</tr>
<tr>
<td>23</td>
<td>Costa Rica</td>
<td>2012</td>
</tr>
<tr>
<td>24</td>
<td>India</td>
<td>2012</td>
</tr>
<tr>
<td>25</td>
<td>Moldova</td>
<td>2012</td>
</tr>
<tr>
<td>26</td>
<td>Nepal</td>
<td>2012</td>
</tr>
<tr>
<td>27</td>
<td>Pacific Islands</td>
<td>2012</td>
</tr>
<tr>
<td>28</td>
<td>Papua New Guinea</td>
<td>2011</td>
</tr>
<tr>
<td>29</td>
<td>Brazil</td>
<td>2011</td>
</tr>
<tr>
<td>30</td>
<td>Lao PDR</td>
<td>2011</td>
</tr>
<tr>
<td>31</td>
<td>Thailand</td>
<td>2011</td>
</tr>
<tr>
<td>32</td>
<td>Malawi</td>
<td>2011</td>
</tr>
<tr>
<td>33</td>
<td>Senegal</td>
<td>2011</td>
</tr>
<tr>
<td>34</td>
<td>Jamaica</td>
<td>2011</td>
</tr>
<tr>
<td>35</td>
<td>El Salvador</td>
<td>2011</td>
</tr>
<tr>
<td>36</td>
<td>Mongolia</td>
<td>2011</td>
</tr>
<tr>
<td>37</td>
<td>Ghana</td>
<td>2011</td>
</tr>
<tr>
<td>38</td>
<td>Bangladesh</td>
<td>2011</td>
</tr>
<tr>
<td>39</td>
<td>Somalia</td>
<td>2010</td>
</tr>
<tr>
<td>40</td>
<td>Georgia</td>
<td>2010</td>
</tr>
<tr>
<td>41</td>
<td>Libya</td>
<td>2010</td>
</tr>
<tr>
<td>42</td>
<td>Cambodia</td>
<td>2010</td>
</tr>
<tr>
<td>43</td>
<td>China</td>
<td>2010</td>
</tr>
<tr>
<td>44</td>
<td>Maldives</td>
<td>2010</td>
</tr>
<tr>
<td>45</td>
<td>Turkey</td>
<td>2010</td>
</tr>
<tr>
<td>46</td>
<td>Indonesia</td>
<td>2010</td>
</tr>
<tr>
<td>47</td>
<td>Guyana</td>
<td>2010</td>
</tr>
<tr>
<td>48</td>
<td>Zambia</td>
<td>2010</td>
</tr>
<tr>
<td>49</td>
<td>Seychelles</td>
<td>2009</td>
</tr>
<tr>
<td>50</td>
<td>Chile</td>
<td>2009</td>
</tr>
<tr>
<td>51</td>
<td>Tajikistan</td>
<td>2009</td>
</tr>
<tr>
<td>52</td>
<td>Republic of South Africa</td>
<td>2009</td>
</tr>
<tr>
<td>53</td>
<td>Uganda</td>
<td>2009</td>
</tr>
<tr>
<td>54</td>
<td>Botswana</td>
<td>2009</td>
</tr>
<tr>
<td>55</td>
<td>Afghanistan</td>
<td>2009</td>
</tr>
<tr>
<td>56</td>
<td>Guatemala</td>
<td>2009</td>
</tr>
<tr>
<td>57</td>
<td>Barbados and OECS</td>
<td>2009</td>
</tr>
<tr>
<td>58</td>
<td>Uzbekistan</td>
<td>2009</td>
</tr>
<tr>
<td>59</td>
<td>Bosnia and Herzegovina</td>
<td>2009</td>
</tr>
<tr>
<td>60</td>
<td>Philippines</td>
<td>2009</td>
</tr>
<tr>
<td>61</td>
<td>Burkina Faso</td>
<td>2009</td>
</tr>
<tr>
<td>62</td>
<td>Peru</td>
<td>2009</td>
</tr>
<tr>
<td>63</td>
<td>Ecuador</td>
<td>2008</td>
</tr>
<tr>
<td>64</td>
<td>Republic of Congo</td>
<td>2008</td>
</tr>
<tr>
<td>65</td>
<td>Benin</td>
<td>2008</td>
</tr>
<tr>
<td>66</td>
<td>Rwanda</td>
<td>2008</td>
</tr>
<tr>
<td>67</td>
<td>Argentina</td>
<td>2008</td>
</tr>
<tr>
<td>68</td>
<td>Bhutan</td>
<td>2007</td>
</tr>
<tr>
<td>69</td>
<td>Nicaragua</td>
<td>2007</td>
</tr>
<tr>
<td>70</td>
<td>Colombia</td>
<td>2007</td>
</tr>
<tr>
<td>71</td>
<td>Jordan</td>
<td>2007</td>
</tr>
<tr>
<td>72</td>
<td>Serbia</td>
<td>2006</td>
</tr>
<tr>
<td>73</td>
<td>Honduras</td>
<td>2006</td>
</tr>
<tr>
<td>74</td>
<td>Ethiopia</td>
<td>2006</td>
</tr>
<tr>
<td>75</td>
<td>Montenegro</td>
<td>2006</td>
</tr>
<tr>
<td>76</td>
<td>Lao PDR</td>
<td>2006</td>
</tr>
<tr>
<td>77</td>
<td>Bangladesh</td>
<td>2005</td>
</tr>
<tr>
<td>78</td>
<td>Syria</td>
<td>2005</td>
</tr>
<tr>
<td>79</td>
<td>China</td>
<td>2005</td>
</tr>
<tr>
<td>80</td>
<td>Turkey</td>
<td>2004</td>
</tr>
<tr>
<td>81</td>
<td>Yemen</td>
<td>2004</td>
</tr>
<tr>
<td>82</td>
<td>Ukraine</td>
<td>2004</td>
</tr>
<tr>
<td>83</td>
<td>Egypt</td>
<td>2004</td>
</tr>
<tr>
<td>84</td>
<td>Nigeria</td>
<td>2004</td>
</tr>
<tr>
<td>85</td>
<td>Jamaica</td>
<td>2004</td>
</tr>
<tr>
<td>86</td>
<td>Mozambique</td>
<td>2004</td>
</tr>
<tr>
<td>87</td>
<td>Vietnam</td>
<td>2003</td>
</tr>
<tr>
<td>88</td>
<td>Bulgaria</td>
<td>2003</td>
</tr>
<tr>
<td>89</td>
<td>Sudan</td>
<td>2002</td>
</tr>
<tr>
<td>90</td>
<td>India</td>
<td>2002</td>
</tr>
</tbody>
</table>
Empowered lives.
Resilient nations.

Independent Evaluation Office
220 East 42nd Street
20th Floor
New York, NY 10017, USA

Tel. +1(646) 781 4200
Fax. +1(646) 781 4213
www.undp.org/evaluation

Connect with us: