



**Executive Board of the
United Nations Development
Programme, the
United Nations Population Fund and
the United Nations Office for Project
Services**

Distr.: General
10 November 2014

Original: English

First regular session 2015

26 - 30 January 2015, New York

Item 4 of the provisional agenda

Evaluation

Review of the UNDP evaluation policy

Contents

<i>Chapter</i>	<i>Page</i>
I. Scope of the review	2
II. Methodology.....	2
III. Key findings: the performance of UNDP and associated entities with responsibilities for implementing evaluation policy against the theory of change.....	3
IV. Major issues and options for the future of the UNDP evaluation policy	6



I. Scope of the review

1. The review assesses the performance of the UNDP evaluation function since 2011, the extent to which the organization has responded to the requirements of the policy, and the status of policy implementation. It identifies areas that require policy changes or management decisions to improve the evaluation function. As stipulated in the terms of reference and confirmed during its inception mission, the review focuses on three elements of the evaluation policy: the overall UNDP evaluation function; independent evaluations; and decentralized evaluations¹. Given reported recurring weaknesses in the coverage and quality of decentralized evaluations and the special interest of the Executive Board in this issue, the review pays particular attention to the third element, confirming existence of weaknesses, exploring the reasons for them, and suggesting appropriate courses of action. For the associated funds and programmes, namely UNCDF and UNV, the review assesses how they have implemented evaluation recommendations and compares that with the specifications of the policy and the approaches adopted by the Independent Evaluation Office.

2. The evaluation policy assigns key functions to different entities in the overall UNDP system. The Executive Board is the custodian of the evaluation policy. The Independent Evaluation Office is the custodian of several aspects of evaluation, with particular responsibility for the delivery of independent evaluations (notably assessments of development results and programme and thematic evaluations). The Administrator is accountable for UNDP results, including responsibility for ensuring that bureaus and country offices, under the oversight of their senior managers, deliver appropriately independent and impartial decentralized evaluations. The Evaluation Units of the associated funds and programmes are custodians of the evaluation function for their parent bodies.

II. Methodology

3. The review took a 'theory of change' approach, using mixed methods and triangulation to ensure data quality. Qualitative and quantitative data relevant to the policy and its implementation were generated through several lines of enquiry, including:

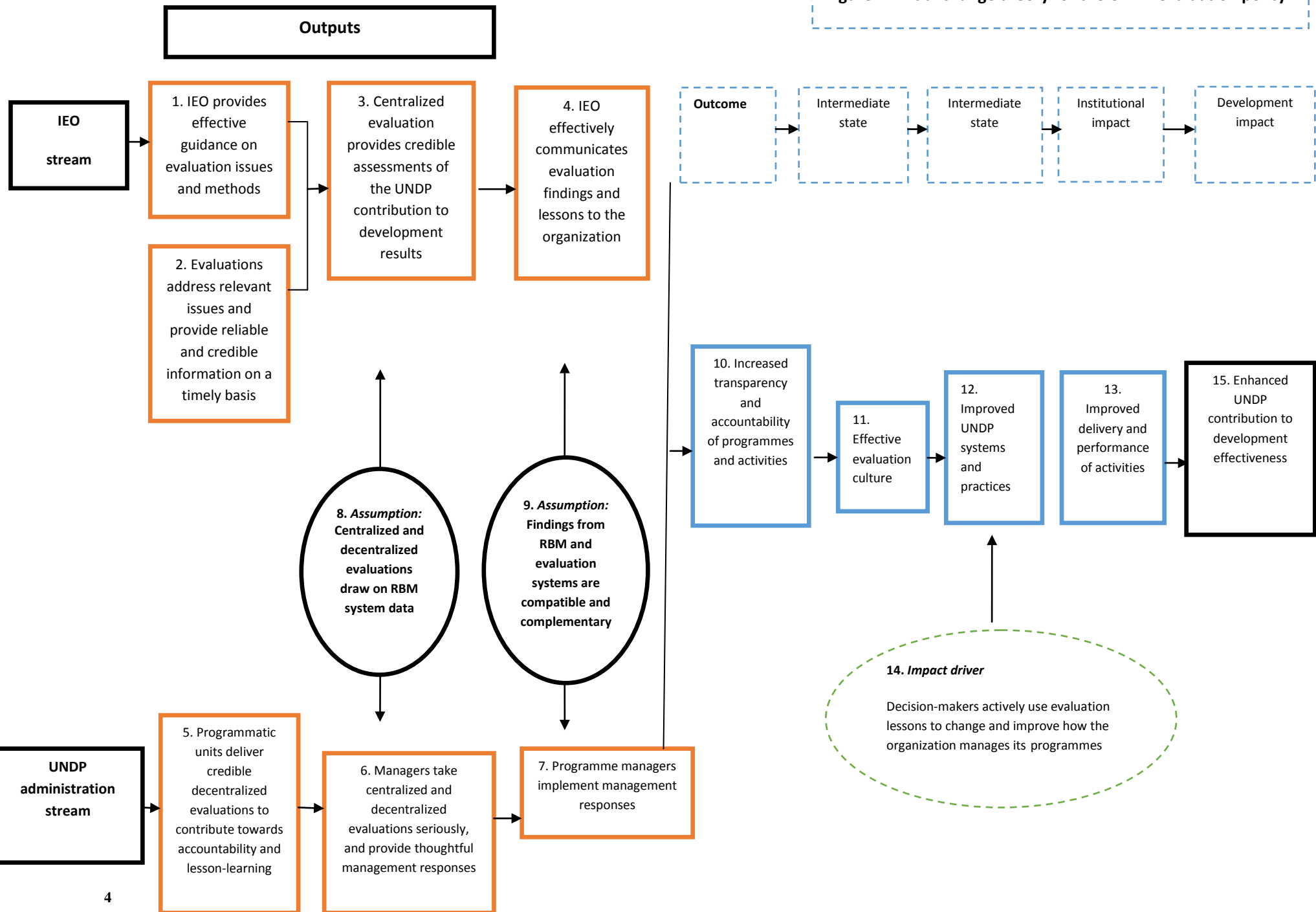
- (a) A document review covering a wide range of documents and files;
- (b) A meta-evaluation (quality assessment) of a sample of decentralized evaluations and evaluations completed by the Independent Evaluation Office since 2012, to review the quality of UNDP evaluations and of the quality assessment process;
- (c) Missions to UNDP headquarters, country offices and regional service centres in Africa, Asia and the Pacific, Europe, Latin America and the Caribbean, and the UNV office in Bonn. The missions included semi-structured interviews or focus groups with a wide range of relevant stakeholders in UNDP, multilateral organizations, bilateral donors and partner country institutions, all of which produced additional documents for review;
- (d) Electronic surveys soliciting views from UNDP resident representatives, staff with evaluation responsibilities, and external consultants who have conducted decentralized evaluations for UNDP over the past three years; and
- (e) Mini case studies of evaluation in UNV and UNCDF based on interviews and document analysis.

¹ Independent evaluations are conducted by the Independent Evaluation Office, while decentralized evaluations are conducted by independent external experts commissioned by UNDP programme units.

III. Key findings: the performance of UNDP and associated entities with responsibilities for implementing evaluation policy against the theory of change

4. The key findings of the review have been analysed within the framework of a theory of change (figure 1). This shows that the policy provides for two main streams of evaluation output to be delivered by the Independent Evaluation Office (boxes 1-4) and UNDP management (boxes 5-7). According to the theory, those streams are interlinked (boxes 8 and 9). The deliverables from the two streams should merge to produce an effective evaluation culture that would drive improved institutional delivery and performance (boxes 10-14). The main driving force in the latter steps of the process would be the active use of verified evaluation lessons on the part of decision-makers to improve how UNDP works.

Figure 1. Initial change theory for the UNDP evaluation policy



5. Table 1, below, provides a detailed rating of progress along the results chain shown in figure 1. It shows that there is minimal movement towards the intended impact of a demonstrably improved UNDP contribution to development effectiveness. This is because the main outcome necessary in order to move towards that impact, namely, “increased transparency and accountability of programmes and activities”, has not been delivered. In fact, given the current challenges to the overall evaluation system, even if UNDP were to enhance its overall development effectiveness, it would lack credible evidence that this was the case. The present review regards this as a very serious situation, which the Executive Board may wish to address urgently through amendments to the evaluation policy and careful monitoring of the results of such amendments.

Table 1. Rating of progress along causal pathway of theory of change for evaluation policy

Element of theory (see figure 1, above)	Rating	Summary of analysis
1. IEO provides effective guidance on evaluation issues and methods.	++	Guidance, mainly given through handbooks and web-based training, is widely used but does not meet all needs of staff throughout the organization, and is in need of updating.
2. Independent evaluations address relevant issues and provide timely, reliable and credible information.	++	Quality of evaluations generally good, but some problems with timely completion.
3. Centralized evaluation provides credible assessments of UNDP contribution to development results.	++	Quality of evaluations generally good, but some problems with credibility of analysis of UNDP contribution, particularly when drawing on outcome information from management.
4. IEO effectively communicates evaluation findings and lessons to the organization	++	Good quality reports but limited range of dissemination products restricts access to information generated.
5. Programmatic units deliver credible decentralized evaluations to contribute towards accountability and lesson-learning	- - -	Over 38% of decentralized evaluation consultants’ report some form of unacceptable measures by UNDP managers affecting the independence of their evaluations. Credibility of information in the ERC is thus severely compromised, reducing its value for accountability or lesson-learning. There is also uneven quality and a preponderance of low-cost evaluations.
6. Managers take centralized and decentralized evaluations seriously and provide thoughtful management responses.	-	Since many decentralized evaluation reports are compromised, management responses are of widely varying quality and utility. Responses to centralized (independent) evaluations are generally more focussed.
7. Programme managers implement management responses	-	At senior management level, concerted measures are in place to monitor implementation of management responses. However, unreliable quality of decentralized evaluations reduces the effectiveness of this process.
8. Assumption: centralized and decentralized evaluations draw on	++	Despite some issues with availability and quality of data in RBM systems (including the ROAR), they are widely used by both centralized and

RBM system data		decentralized evaluations.
9. Assumption: findings from RBM and evaluation systems are compatible and complementary.	+	Findings are broadly compatible, but many decentralized evaluation reports have been amended to present results desired by management (particularly at the country office level).
10. Increased transparency and accountability of programmes and activities.	---	Major flaws in decentralized evaluation system have severely reduced transparency and accountability.
11. Effective evaluation culture.	--	Owing to flawed decentralized evaluation system and excessive opposition to a few independent evaluations, there has been limited progress towards an overall evaluation culture.
12. Improved UNDP systems and practices	-	Although senior management has made substantial efforts to develop improved use of evaluation findings, it has been ineffective in ensuring the veracity of the data upon which its analyses are based.
13. Improved delivery and performance of activities	-	There is insufficient impartial evaluation evidence to verify any reported improvements.
14. Enhanced UNDP contribution to development	-	There is insufficient impartial evaluation evidence to verify any enhanced contribution.
15. Impact driver: decision-makers actively use evaluation lessons to change and improve how the organization manages its programmes.	-	Although there are decision-makers who attempt to use lessons to improve programmes, the lack of veracity in the decentralized evaluation system prevents this from occurring throughout UNDP.

IEO = Independent Evaluation Office; ERC = Evaluation Resource Centre; RBM = results-based management; ROAR = results-oriented annual report

IV. Major issues and options for the future of the UNDP evaluation policy

6. The findings and analysis presented in this report show that there are several issues that the Board may wish to address through a revision of the evaluation policy. The review recommends that the revised policy should provide means through which to overcome all of the identified challenges. Those challenges are of different magnitudes, however, and will require various levels of input and Board attention. This chapter, therefore, devotes most of its attention to issues surrounding decentralized evaluations, which are currently the biggest element preventing UNDP from developing and benefiting from an effective evaluation culture. Unless these issues are overcome, the Board will continue to lack a coherent set of independent and impartial evidence on field level performance upon which to base its decisions.

7. The review presents initial (sometimes alternative) approaches to address the issues raised. It recommends that these should be advanced and made specific by detailed discussions among the Board members, and between the Board, UNDP management and the Independent Evaluation Office.

Issue 1. Overall reliability of decentralized evaluations severely compromised

The challenge

8. The policy places the onus for ensuring the quality of decentralized evaluations on the UNDP Administrator and senior management, with a limited 'quality assessment' role for the Independent Evaluation Office. This approach has not worked. The symptoms of this failure are as follows:

- (a) There is a substantial rate² of interference by management (mainly at the country office level) in evaluation reports of 'independent' consultants;
- (b) The quality assessment system therefore rates reports of unknown but varying veracity;
- (c) The Evaluation Resource Centre compiles information on and from unreliable reports, which is therefore inaccurate;
- (d) Much of the analysis and reporting on decentralized evaluations in the Evaluation Resource Centre system is thus based on inaccurate or unreliable data; and
- (e) Management reporting on decentralized evaluations to the Executive Board is based on aggregated data from a seriously flawed database, which is not suitable for helping the Board to provide direction to UNDP.

9. To understand how this situation might best be changed, it is necessary to determine which factors have contributed towards it, either as a cause or as part of the institutional environment that has allowed it to develop. These can be summarized as follows:

- (a) Final evaluations are often not seen as useful, particularly by country offices;
- (b) Few staff dealing with evaluation have any specialized training in the subject;
- (c) There is a high reliance by staff on one handbook, which is useful, but incomplete and out of date;
- (d) The concept of independence of decentralized evaluations is often not understood or supported;
- (e) Evaluations are seen as undermining funding efforts;
- (f) Managers believe that some consultants are biased against UNDP or do not understand it;
- (g) Resources are often inadequate for high-quality evaluations;
- (h) Evaluation is not a priority for many national counterparts;
- (i) Many managers have not prioritized the development of an evaluation culture (although some have);
- (j) Policy makes UNDP management bodies responsible for quality assurance of decentralized evaluations, but this has not been implemented as intended: there is only quality assessment of reports by the Independent Evaluation Office;
- (k) No effective mechanism exists to resolve disputes between consultants and clients; and
- (l) Prior experience of disputes, combined with inadequate resources to meet the requirements of the terms of reference, deter many consultants from bidding for UNDP decentralized evaluation work, leading to a restricted choice of consultants.

10. Underlying the above array of contributory factors is the perception among staff at various levels that neither "sticks nor carrots" are in place to encourage desired management behaviour towards independent evaluations (or indeed any evaluation). Each manager is free to form a personal approach towards evaluation, since performance assessments rarely reward positive performance in this area or penalize negative aspects. It is usually only the relatively junior 'monitoring and evaluation specialists' who are judged on this aspect, and they are not sufficiently senior to make a major difference to the overall approach of the offices they serve.

² About 38 per cent of consultants who responded to a survey (254 respondents; response rate of 50.8% of the population; margin of error 3.6 per cent) reported some form of unacceptable interference by management.

Recommended approach³

11. *The evaluation policy should require management to introduce and enforce effective quality assurance systems for decentralized evaluations; with verification by the Independent Evaluation Office and penalties for units⁴ that do not comply with standards. This should be supported by updated and additional guidelines and an assessment by the Executive Board of the value added to available management information by large numbers of low-budget evaluations of variable quality.*

12. At present, the directors of the regional bureaus are responsible for “ensuring quality and implementation of evaluation (plans and) practices”. However, those bureaus have gradually reduced the ‘monitoring and evaluation’ staff posts that would enable them to fulfil this function, even though it is required of them. Furthermore, documents available to the independent review suggest that, under the new UNDP structural review, even less staff time may be devoted to evaluation than is presently the case. There are approximately 300 decentralized evaluations per year requiring effective quality assurance by the five regions. Given that “quality assurance” would need to begin with such issues as the quality of terms of reference and evaluability analysis, and follow the process right up to the final report, the review concludes that UNDP management could not deliver this with less than 10 full-time equivalent evaluation advisers⁵ on site across the regions. They could be either regional staff or jointly funded by the Independent Evaluation Office and the regions (assuming additional resources were made available). Matrix management would need to be applied, with prime responsibility falling to the Independent Evaluation Office but with some level of operational management supplied by the regions.

13. If the necessary evaluation advisers were in place, the next step would be independent verification of evaluation reports. Two measures would be necessary here, one of which would require further resources. The first would parallel a process adopted by the UNDP-Global Environment Facility office for terminal evaluations. It would require that independent consultants submit their draft reports simultaneously to the commissioning unit (usually the country office) and to the regional evaluation adviser. This would enable the adviser to verify subsequently that no surprising changes occurred between the draft and the final version. Such changes would trigger follow-up enquiries by the adviser. The second measure would require a follow-up evaluation mission by (or on behalf of) the Independent Evaluation Office to a percentage (the review suggests 10 per cent) of projects. This would verify the quality of the report by reassessing the project, including its ratings⁶. Since the projects to be verified would not be notified until after the original decentralized evaluation had been submitted, this would introduce some “teeth” into the assessment process, which could be strengthened by ensuring comparable verification coverage across regions and issuing “red flags” to regional and country offices that were found to have submitted inappropriately influenced or amended evaluation reports.⁷

³ The final report of the review offers alternative options, but the present executive summary presents only the approach that the review recommends.

⁴ A way of introducing penalties into the system practiced among some international finance institutions is to give “red flags” to management units breaching regulations or standards. These can accumulate and be considered in future performance assessments.

⁵ ‘Monitoring and evaluation specialists’ could be counted towards the required number of evaluation advisers according to whether they are full- or part-time in their monitoring and evaluation capacity, and based on the percentage of their time devoted to evaluation rather than monitoring.

⁶ This would follow procedures similar to some of those used in the World Bank project performance assessment system. Although the Bank conducts project performance assessment reports in 20-25 per cent of projects previously assessed by an implementation completion report and an implementation completion report review (a desk study), the present review proposes that the UNDP process assess only 10 per cent of projects. This is justified by the fact that UNDP decentralized evaluations are intended to be conducted by independent experts, whereas World Bank implementation completion reports may be conducted by project implementers.

⁷ The Office of Audit and Investigations is already able to raise “audit issues” concerning specific offices, which the offices subsequently have to address.

14. Since there are around 300 decentralized evaluations per year, this would require roughly 30 follow-up missions each year. As these missions are primarily for verification, they could be shorter than the original evaluation, but this total would still require additional funding.

15. In support of the above control measures, the Independent Evaluation Office and UNDP need to collaborate to update and expand the *Handbook on Planning, Monitoring and Evaluating for Development Results*, particularly in such areas as gender equity, creating and maintaining evaluation partnerships, joint evaluation, protecting independence, and evaluation ethics for managers. The handbook should be supplemented by enhanced opportunities for evaluation training, both on-line and in person. The Independent Evaluation Office should also carefully examine the numerous guidance products developed by the United Nations Evaluation Group, with a view to preparing and circulating concise applications of their messages for UNDP.

16. The review suggests that, if these measures were all put in place, including with adequate staff actually in post (not “positions”) it would enable UNDP over time to develop a credible decentralized evaluation system, which is currently not in place. This would provide both management and the Board with the essential body of verified data on country level performance to effectively perform their respective roles.

Issue 2. Unreasonable disputes over some independent evaluations

The challenge

17. The majority of independent evaluations are delivered without major problems. However, several (five out of 22 assessments of development results in the last three years) have become embroiled in unreasonable disputes, which have resulted in delays of up to a year in the finalization of reports. The review indicates that three of these originated from a country office and eventually involved the relevant regional bureau, while two are attributed to delays caused by Independent Evaluation Office processes.

18. The symptoms of this issue are:

- (a) Time scales are disrupted and delayed;
- (b) Evaluations are not delivered as planned for management purposes;
- (c) Independent Evaluation Office resources are diverted from other intended work;
- (d) The value added from independent evaluations is reduced by the unwillingness of some managers to encompass external perspectives on their work; and
- (e) There is reputational risk, since disputes become public knowledge.

19. Factors contributing to the challenge are:

- (a) Managers perceive the quality of independent evaluations as variable (which could be true)
- (b) Managers believe that some consultants are biased against UNDP or do not understand it, (which could be true);
- (c) Evaluations are less positive than management expects, and are seen as undermining “communication strategies”;
- (d) The concept of independence of centralized evaluations is interpreted differently by stakeholders in different positions; and
- (e) Presentation of “negative” findings is perceived as disloyal by some managers.

20. The Independent Evaluation Office has accepted that it has contributed to this situation and has introduced measures to strengthen its approach. These include a rebalancing of the role of Independent Evaluation Office professional staff and consultants, with a stronger emphasis on its direct management of such evaluations. Also, it has introduced procedures to expedite progress, through enforcement of timelines.

Recommended approach

21. *In the absence of timely feedback from management parties, the Independent Evaluation Office should be given clear authority to proceed with all steps in the independent evaluation process according to the pre-established timetable. Enduring objections by management to evaluation findings or recommendations should be reserved for inclusion in the management response and should not be allowed to delay national workshops or the publication of reports. UNDP units (such as country offices) that hinder the progress or completion of independent evaluations should be penalized through a “red flag” system.*

22. This approach could eliminate lengthy disputes and delays with independent evaluations at no additional cost. A potential disadvantage is that national partners (or other parties affected by thematic or global evaluations) may get conflicting messages if they receive reports from the Independent Evaluation Office that have not been accepted by UNDP management. On the other hand, this could become an advantage, since it could place pressure on management to respond promptly and constructively, enabling evaluations to be completed on time. That pressure would be reinforced by the issuance of red flags to managers who are seen to obstruct the system. The review received ample evidence from UNDP staff and managers at different levels that one of the key factors seen as reducing the effectiveness of the evaluation function is that the system lacks penalties against those who behave inappropriately.

Issue 3. Policy too vague on issues essential to the operational independence of the Independent Evaluation Office

The challenge

23. The policy is vaguely worded or silent on a number of issues that could together pose serious challenges to the operational independence of the Independent Evaluation Office. The symptoms include:

- (a) The appointment term of the Director of the Independent Evaluation Office is “up to” four years, which allows management the possibility of offering less;
- (b) The duration of a possible renewal is not specified;
- (c) The Executive Board does not chair or even sit on the appointment committee for the Director or Deputy Director, while management has the majority;
- (d) It is not clear how the Independent Evaluation Office Director’s performance is assessed, while such clarity is essential for quality assurance and confidence in the position; and
- (e) The policy does not specify that the Director of the Independent Evaluation Office can submit reports directly to the Executive Board as necessary in the event of serious disputes with management.

24. Factors contributing to this situation appear to include:

- (a) Assumptions of a harmonious relationship between the Independent Evaluation Office and UNDP management have not been realized;
- (b) The implications of vague wording have not been fully appreciated.
- (c) The importance of the credibility of the Independent Evaluation Office Director, through appropriate performance assessment, has not been emphasized.

Recommended approach

25. *The review recommends that the Executive Board amend the policy to specify the lead role of the Board in recruitment procedures for the Director of the Independent Evaluation Office; the duration of the post (subject to performance assessment); renewal processes and duration; and the power of the Director to report directly to the Board as necessary. These measures would strengthen the structural independence of the Office, in keeping with its new title.*

Issue 4. No assurance of independent evaluation by associated funds and programmes

The challenge

26. The policy describes the Independent Evaluation Office as the “custodian of the evaluation function” for independent evaluations in UNDP, specifying that the “evaluation units of the associated funds and programmes are the custodians of the evaluation function in their organizations”. This indicates that the roles of the different evaluation entities are parallel in some respects. However, while the policy states that Independent Evaluation Office “conducts independent evaluations in line with best international evaluation standards”, the evaluation units simply “manage and conduct evaluations”. The review concludes that, although the evaluation units cannot have the same degree of structural independence as the Independent Evaluation Office because of the small management systems of the institutions, they should nevertheless be required to operate in line with “best international evaluation standards”.

27. Symptoms of this challenge include:

- (a) Owing to the small size of associated fund bodies, evaluation units cannot be structurally independent of management;
- (b) The Independent Evaluation Office and associated fund evaluation units are all defined as “custodians of the evaluation function”;
- (c) Evaluation units therefore have parallel status to the Independent Evaluation Office. Their independent evaluations are self-verified, but at a lower level than is the case for the Independent Evaluation Office (P4 as against D2);
- (d) The small size of units means that all their evaluations are undertaken by consultants, so that they are effectively managing decentralized evaluations; and
- (e) The independence and impartiality of evaluations conducted by the associated funds and programmes (although currently acceptable) is not guaranteed by the policy.

28. Factors contributing to the situation include:

- (a) Structural challenges to independence in small organizations not incorporated in the policy;
- (b) Levels of evaluation expertise and the authority of the chief are set at a lower level for associated funds and programmes than for the Independent Evaluation Office, with no compensating quality assurance; and
- (c) The level of collaboration between the Independent Evaluation Office and the evaluation units of the associated funds and programmes is inadequate to develop consistent approaches to quality.

Recommended approach

29. *The review notes that the evaluation units of the associated funds and programmes have developed in quite different ways, and recommends that the policy should incorporate new and flexible approaches as follows: **evaluation units should be required to either submit their independent evaluations to the Independent Evaluation Office for quality assurance (to ensure comparable “best international evaluation standards”); or to collaborate directly with the Independent Evaluation Office to manage and report on their independent evaluations. Associated with these measures, it is also recommended that the policy should require the Independent Evaluation Office to pay more systematic attention to the contribution of the associated funds and programmes to UNDP results in all of its independent evaluations.***

30. The review considers the proposed quality assurance to be appropriate because it is fundamentally the director-level posts of the Independent Evaluation Office that enable it to operate at the level of best international evaluation standards, while appointments at this level are unlikely to be viable in the evaluation units of small associated funds and programmes. The alternative is seen as a form of management collaboration on evaluations, which could be formally agreed between a

fund and the Independent Evaluation Office and subject to periodic review. Such collaboration could, for example, include a fund or programme's financing a position (full- or part-time) in the Independent Evaluation Office focusing on its evaluation work, or simply commissioning the office to conduct (or conduct jointly) major evaluations for the fund or programme concerned.

31. The review concludes that these measures would validate the independence of evaluations from the units and ensure comparability of standards across the evaluation entities covered by the policy.

Issue 5. Relevance of concepts in the policy needs updating

The challenge

32. The policy contains a section on definitions (paragraphs 5 and 6) that the review finds problematical. The definitions are mainly derived from the glossary of key terms in evaluation and results-based management published in 2000 by the Organisation for Economic Co-operation and Development. In the light of discussions and debates within the evaluation community over the last decade, some of the definitions do not fully accord with current thinking, while other important topics are not mentioned.

Recommended approach

33. *The review recommends that the section on definitions in the policy be replaced by a more general text indicating that the Independent Evaluation Office will periodically update and disseminate current evaluation topics and definitions on the basis of best international standards through operational handbooks and other appropriate means.*
