

Independent Review of UNDP Evaluation Policy

18 MARCH 2019

Table of Contents

1. Introduction	2
1.1. Background.....	2
1.2. Purpose, Scope and Objectives	2
1.3. Changes in the Context	2
2. Assessment of the evaluation policy and its implementation	3
2.1. Strengths of the evaluation policy and its implementation.....	5
2.2. Weaknesses of the evaluation policy and its implementation	6
2.3. Discussion of the key constraints related to the evaluation policy and its implementation.....	7
3. Assessment of the Evaluation Architecture	8
3.1. Strengths of the Evaluation Architecture.....	8
3.2. Weaknesses of the evaluation architecture.....	9
3.3. Discussion of the key constraints related to the evaluation architecture	10
4. Conclusions and Recommendations	12
4.1. Conclusions	12
4.2. Recommendations	12
5. Annex 1.....	14
List of Persons Interviewed	14
6. Annex 2.....	16
Methods	16
7. Annex 3.....	22
Terms of Reference	22

List of Figures and Tables

Figure 1: A simplified theory of change	3
Figure 2: Evaluation policy trade-offs	4
Figure 3: Emerging themes from a qualitative analysis of the interview transcripts	5
Figure 4: Emergent themes, phase I	17
Figure 5: Coding co-occurrences and similarity index, phase II	17
Figure 6: Emergent themes, phase III (final)	18
Table 1: Quality of decentralized evaluations in 2016 and 2017	6
Table 2: Number of regional M&E specialists	9

1. Introduction

1.1. Background

The first UNDP Evaluation Policy was approved by the UNDP Executive Board (EB) in 2006, and subsequently revised in 2011. Following an independent review of the Policy in 2014, the next and current iteration of the UNDP Evaluation Policy was approved by the Board in September 2016. In its decision approving the 2016 Policy, the Board requested that it be reviewed in 2019, with a report to the Board at its annual Session in June 2019.

The review is based on interviews with a cross-section of stakeholders and a desk review of key documents. It is done within a timeframe of approximately 40 working days. The review team consisted of Osvaldo Feinstein, team leader (member of UNDP/IEO Evaluation Advisory Panel), Patty Chang (Adjunct Associate Professor, NYU Center for Global Affairs), and Per Øyvind Bastøe (Evaluation Director, Norad and DAC EvalNet Chair).

1.2. Purpose, Scope and Objectives

The review covers the period from September 2016 to January 2019, taking into consideration contextual changes since the approval of the 2016 Evaluation Policy. The review encompasses a set of key aspects of the evaluation policy and is expected to:¹

- Assess the progress made in implementing the revised evaluation policy of 2016, noting the strengths and weaknesses;
- Review the evaluation architecture; and
- Identify constraints inhibiting the effective implementation of the policy and areas that may require policy change or management decision to improve the evaluation function.

The findings and recommendations will be presented to the UNDP Executive Board during the annual session of the Executive Board in June 2019.

1.3. Changes in the Context

Since the approval in September 2016 of the Evaluation Policy there were several important changes in the context in which the new policy was, and continues to be, implemented:

- a) The appointment in June 2017 of a new UNDP Administrator who fully understands the role of evaluation and its potential contribution to learning;
- b) The Sustainable Development Goals (SDGs) and the 2030 Agenda;
- c) The UN Reform process, with the repositioning of the UN development system;
- d) The guidelines for United Nations Development Assistance Frameworks (UNDAFs) and mandatory requirements for UNDAF evaluations introduced in 2017; the new UNDAF guidance will be completed by March 2019. UNDP is committed to making UNDAF the primary planning instrument;
- e) The UNDP Strategic Plan, 2018-2021;

¹ Terms of Reference, Independent Review of the UNDP Evaluation Policy, January 2019.

- f) The System-wide Evaluation in Support of UN Reform Initiative;
- g) The Charter of the Independent Evaluation Office of UNDP, published in December 2018; and
- h) UNDP Evaluation Guidelines, published in January 2019.

Thus, UNDP’s Evaluation Policy has been implemented in a dynamic context. These developments will be considered in the following sections of this Review.

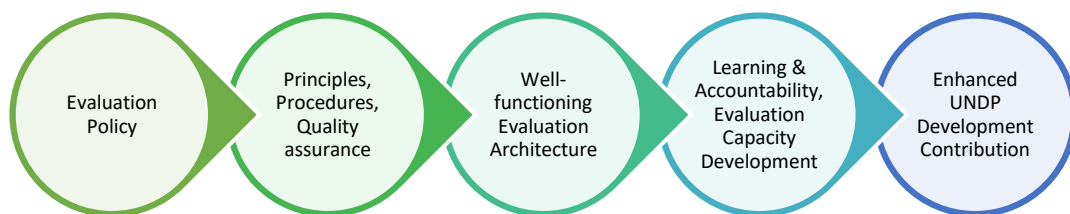
2. Assessment of the evaluation policy and its implementation

The UNDP evaluation policy sets out the purpose and basic principles of evaluation for UNDP and its associated funds and programmes.²

The Policy defines the institutional architecture and clarifies the roles and responsibilities within the UNDP institutional framework, providing a policy foundation for safeguarding the independence of evaluations. The Policy covers all evaluations conducted by the Independent Evaluation Office of UNDP, as well as those commissioned by UNCDF, UNV, Programme and Policy units (decentralized evaluations). The policy harmonizes the oversight of UNDP’s audit and evaluation functions under an expanded Audit and Evaluation Advisory Committee (AEAC). The Evaluation Policy establishes financial benchmarks for the UNDP evaluation function, stipulating that subject to availability of resources, it would allocate 1% of combined programmatic (core and non-core) resources for the evaluation function annually, within which 0.2% is for the work of the Independent Evaluation Office of UNDP.

Before presenting the findings regarding strengths and weaknesses in the policy and its implementation, it is worthwhile to present a simple framework that attempts to show the expected processes and results of the evaluation policy. Figure 1 is a simplified theory of change: it starts with the evaluation policy, designed with the use of generally accepted evaluation principles adapted to the organizational reality of UNDP.

Figure 1: A simplified theory of change

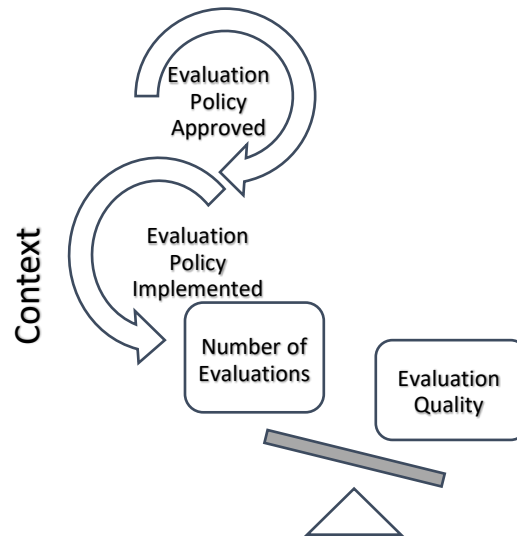


This simplified theory of change can be complemented with the following diagram (Figure 2) which shows that the evaluation policy approved by the EB is implemented in a specific context, yielding a number of evaluations and also having an effect on their

² UN Volunteers (UNV), and the UN Capital Development Fund (UNCDF).

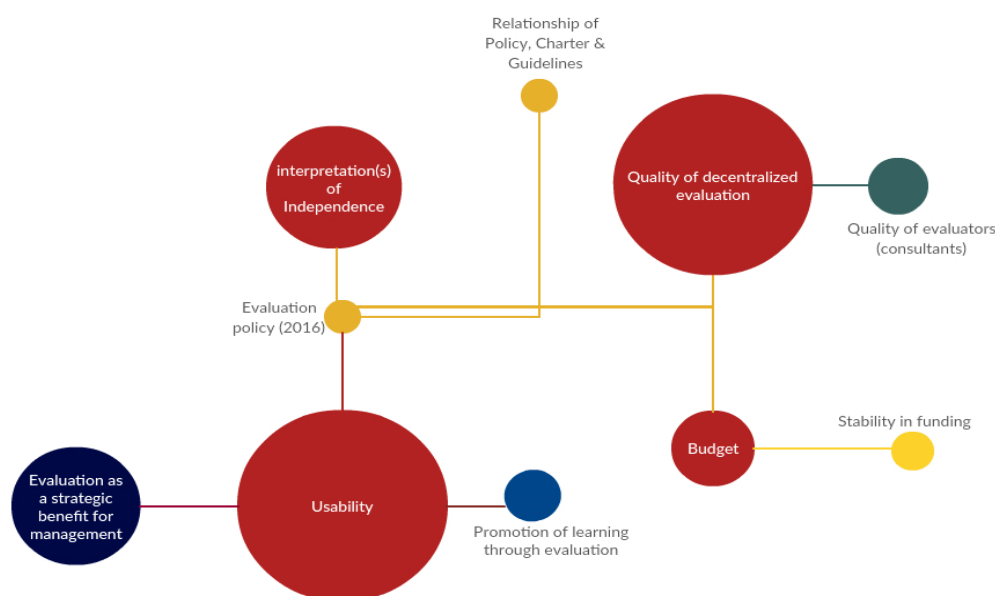
quality, with possible trade-offs between quantity and quality of evaluations. Whereas the theory of change focuses on processes and intermediate and final outcomes, Figure 2 sheds light on a fundamental aspect concerning outputs: the possibility that the expansion in the number of evaluations may jeopardize their quality. Another possible issue, not shown but which may be important, is the trade-off between compliance (with management response) and learning (from evaluations), which may be intensified with an increase in the number of evaluations and the requisite to provide management response to all evaluations.

Figure 2: Evaluation policy trade-offs



The findings on the strengths and weaknesses in the policy and its implementation in this review are based in part on the emerging themes derived from a qualitative analysis of the interview transcripts (see: Annex 2 for the methods). Figure 3 depicts a simplified version of the overarching themes which emerged from the interviews based on a count of thematic frequency (See: Figure 5 in Annex 2 for the full map). The circles are scaled according to a frequency count of the coded themes. The red circles form the highest frequency in mentions, followed by the other colored circles as major sub-themes.

Figure 3: Emerging themes from a qualitative analysis of the interview transcripts



2.1. Strengths of the evaluation policy and its implementation

Clarity: In the interviews the *clarity* of the policy was praised, although some concern was expressed on the lack of precision with respect to the budget for evaluation.

Independence: The way in which independence is framed in the evaluation policy was considered appropriate and a source of credibility of the evaluations produced by IEO. The establishment of the independence of IEO's Director and of the Evaluation Office has been an important factor for enhancing the *credibility* of the evaluations, which also influences the use of evaluations.

Evaluation Use and Learning: Despite the lack of explicit references to evaluation *use* in the policy, it was widely acknowledged that IEO evaluations are used to inform the design of UNDP programmers, contributing to *learning* and to the reputation of UNDP as a transparent and learning organization.

Oversight, Support & Quality Assurance for UNCDF and UNV: The arrangements envisaged in the Evaluation Policy concerning UNCDF and UNV are functioning well. In the case of UNCDF their evaluations are a highly managed process and IEO rated highly the quality of UNCDF evaluations. It is also worthwhile to mention that UNV has access to a significant pool of evaluators. Neither of these two organizations suffer from quality issues or capacity problems in their evaluation function (which, as shown in the next subsection, are problems affecting most of UNDP decentralized evaluations). Also, it is worth noting that UNCDF and UNV conduct a very limited number of evaluations, which may have a positive influence in the quality of evaluations, as pointed out in section 2 and illustrated in fig.2.

2.2. Weaknesses of the evaluation policy and its implementation

Decentralized Evaluations: A key weakness that was highlighted several times in the interviews is the low quality of decentralized evaluations. This issue was already mentioned in the 2014 review of the evaluation policy (which in fact was a review of the evaluation function). The data confirms this judgement as shown in Table 1. It is remarkable that there is small variability among regions, that in all of them more than 75% of the evaluations have been assessed as neither satisfactory nor highly satisfactory (in fact this percentage is more than 80% in four of the five regions). It is also worthwhile to observe that the decentralized evaluations of higher quality are those not managed by Country Offices.

Table 1: Quality of decentralized evaluations in 2016 and 2017

Region	Nº of Evaluations	% of highly satisfactory or satisfactory
Africa	88	19
Arab States	24	17
Asia and the Pacific	38	19
Europe and the CIS	55	22
Latin America and the Caribbean	44	18
Global *	12	58
Total	261	21

* Mainly UNDP BPPS, including UNCDF & UNV evaluations.³

Source: IEO (2018), latest available data, elaborated by the Review Team

The Evaluation Guidelines issued in January 2019 addresses some of the factors that may be root causes of the low quality of decentralized evaluations. There are several issues but two are critical and interrelated:

Conflict of interest: Decentralized evaluations are contracted out to external consultants by the Country Office. This arrangement can foster a perverse incentive for external evaluators to produce positive evaluation results to increase their chances of being rehired by the Country Office for future evaluations. The structural arrangement can promote a bias in the selection of consultants. An alternative scenario is when the Country Office interferes with (critical) findings presented in an evaluation report contracted by an external consultant. There is a gap in oversight, as IEO is far removed from the decentralized evaluation process and there is limited capacity on the part of

³ Since the relaunch of the QA system, 10 UNCDF evaluations have been assessed as follows: 3 as 'Highly Satisfactory'; 6 as 'Satisfactory' and 1 as 'Moderately Satisfactory.'

regional hubs and the regional bureaux to systematically oversee the process. A further layer of complexity entails the issue of impartiality due to the operational duties of M&E staff in large Country Offices, who often focus only a certain percentage of their time on evaluations and work on the projects to be evaluated. It then becomes difficult to ascertain if the final evaluation report findings are a result of Country Office interference or an objective assessment on the part of the external evaluator. Although the actual situation would require more evidence, what is clear is that there is a perception of conflict of interest. Interviews with staff from regional bureaux and regional advisors not only raise these concerns, but also there appears to be a wide variation in the qualifications, expertise, and access to a pre-vetted roster of evaluators based on the regions. It should be noted that the Evaluation Policy (paragraph 37) states that “UNDP management shall take all necessary actions to ensure the objectivity and impartiality of the process and persons hired”.

Limited flexibility: Some interviewees considered that the implementation of the evaluation policy was not flexible enough in regions where the programmes did not often fit the mold of development programming and especially in crises or post-conflict settings, which also required more innovative approaches for data collection

2.3. Discussion of the key constraints related to the evaluation policy and its implementation

The key constraints for an appropriate implementation of UNDP’s Evaluation Policy can be reduced to a capacity and a funding constraint, and both are interrelated. Thus, to improve the quality of decentralized evaluations additional funding may be required to invest in evaluation capacity at the regional level, including to enhance the capacity of regional evaluation advisors. It is worthwhile to mention that the Swiss Development Agency has already provided a valuable financial contribution to enhance the quality of decentralized evaluation. There is scope for evaluation partnerships with other bilateral development agencies with experience, expertise and interest in decentralized evaluation.

Strategic use and learning: While IEO have been consistently praised for the quality, credibility and utility of their reports by all persons interviewed, many of those interviewed still bring up how evaluations can be used more strategically with a view to contribute more to the latest Strategic Plan, the repositioning of UNDP in the UN Development System, and to enhance learning on innovations and scaling up. There was also an interest in evaluations to provide further contextualization in its analyses, and that by tailoring more the messages and language for different audiences it would be possible to encourage uptake and promote use. Overall, the sense is that while evaluations are appreciated, the current approaches, methods, and analyses ultimately have limited strategic application for UNDP.⁴

Evaluation Production and Absorption Capacity: A potential weakness mentioned in the interviews is related to the very strong increase in the number of planned evaluations

⁴ Peer review of UNDP’s Evaluation Office on methodology and knowledge sharing (15 January 2013), pg. 6, para. 20. Already mentioned this issue.

for 2018 compared with the number of evaluations conducted in 2017, not only in terms of the capacity to produce quality and timely evaluations⁵ but also with respect to Management's absorptive capacity. References were made (and confirmed by a review of the figures) both to the decision to have 100% coverage of ICPEs, and to the large number of planned decentralized evaluations in comparison with the number of evaluations completed in 2017. For an additional elaboration of this point and its implications see below the first two paragraphs of section 3.3.

Funding for the evaluation function: Several individuals interviewed expressed concern that the language should be further clarified to ensure that 1% of UNDP's budget will be allocated to evaluations. Others were unclear what the 0.2% funds for independent evaluations covered and if the 0.8% funds for decentralized evaluations was adequate and inclusive of UNCDF and UNV. Some interviewees said that the benchmark could be more of an "aspirational target". Finally, some staff argued that funding for central evaluations should be expanded to enable the capacity for IEO to strengthen oversight of decentralized evaluations.

Interpretation of independence: The majority of individuals interviewed praised the strong independence of IEO as a positive aspect (especially in connection with the quality of the evaluations produced), however some senior management brought up to varying degrees that the interpretation of independence by IEO can be at times excessive, which can create unnecessary friction and limit trust.

Confusion regarding the understanding of terms: The interviews revealed that the key terms in the policy (independence, credibility and utility) were understood differently among the interviewees. A clear example is that "independent" is commonly understood as equivalent to "external".

3. Assessment of the Evaluation Architecture

UNDP's evaluation architecture, as envisaged in the 2016 Evaluation Policy, is a multi-level system. The policy describes the roles and responsibilities related to evaluations of the Administrator, UNDP programme and policy units, UNCDF and UNV, IEO and the AEAC.

3.1. Strengths of the Evaluation Architecture

Staffing of IEO: Between 2017 to 2019, IEO experienced an increase in staff capacity with the addition of 10 new staff members to strengthen the office. The current IEO staff capacity as of 2019 consists of 33 staff members (24 Professional staff, 9 General Service).⁶ The expansion enabled IEO to form three main clusters: (1) Independent Country Programme Evaluation Section; (2) Corporate Evaluation Section; and (3) Evaluation Capacity Development Section, along with the Directorate and Operations Section. According to the 2019 IEO organigram, the Country Programme Evaluation

⁵ The new approach to UNDAF in the short run imposes an additional pressure to deliver in time for the UNDAF country programme, and in the medium term may have even a significant effect on UNDP country programmes.

⁶ Based on a review of 2016-2019 IEO Organigrams and confirmed by IEO Operations Section as at 20 February 2019.

Section is the largest cluster with 9 evaluators and one Head of Section, relative to the other two clusters, each consisting of 5-4 evaluators and one Head of Section.

The current staffing level corresponds to IEO’s commitment towards conducting a greater number of ICPEs, providing ad hoc (rather than systematic) support for decentralized evaluations, and an increase in thematic and corporate evaluations.

Evaluation guidelines: In January 2019, IEO published Evaluation Guidelines after a lengthy participatory consultation process with Regional Bureaux, Country Offices and BPPS M&E specialists, with the expectation that such a guidance will improve the quality of these types of evaluations. In 2016, IEO also reinstated a revised quality assessment process of decentralized evaluations after a two-year hiatus, as well as overhauled the Evaluation Resource Centre (ERC) website. The IEO collaborated with the regional bureaux and with BPPS Effectiveness group to conduct regional workshops to encourage regional evaluation support staff and country office M&E Focal Points to discuss M&E challenges, guidance requirements and training needs.

Capacity development: In addition to supporting UNDP’s internal evaluation capacity development, as indicated in the previous paragraph, in line with the third objective of the evaluation policy IEO has contributed to develop evaluation capacities in member countries through National Evaluation Capacities Conferences held in the different regions in which UNDP operates.

3.2. Weaknesses of the evaluation architecture

Limited evaluation capacity of the regional bureaux/hubs: The capacity in the regional hubs and bureaux is too limited to provide effective support and to exercise oversight of the country offices. Table 2 shows the downward trend in the number of regional M&E specialists. While it is difficult to ascertain if M&E specialists have balanced expertise in both monitoring and evaluation, interviews with regional bureaux and staff suggest that they may have more monitoring than evaluation capacity. It should be noted that UNICEF has regional evaluation advisers, with no monitoring responsibilities.⁷

Table 2: Number of regional M&E specialists

Year	2014	2015	2016	2017
Nº	14	13	12	10

Source: IEO (2018) elaborated by the Review team

Cumbersome management response system (MRS): Although the MRS is a mechanism that forces management to pay attention to evaluations, it has become a time-consuming practice. Based on a desk review of the status of recommendation implementation and in interviews carried out for this review, there appears to be significant gaps in management response tracking and particularly in reporting. The

⁷ As an example, the establishing of regional evaluation advisors at P-5 level with technical reporting lines to the centralized evaluation office, was recommended in the 2017 Peer Review of the UNICEF evaluation function as a measure to safeguard the professionalism, impartiality, effectiveness and efficiency of decentralized evaluations.

challenge is two-fold. First between 2016 to 2019, IEO produced an increasing number of evaluations, with some delivered timelier than others. The time slippage not only creates tension between IEO and Senior Management, but also a risk that the rush to produce management responses for all evaluations becomes a mechanical (as opposed to a learning) exercise. Second, concerning utility, some individual interviewed expressed the desire to see substantive evidence of how evaluations influenced UNDP policies and practices outside of the management response tracking system.

Negative value added for evaluation of the AEAC: This is a weakness that affects independent evaluation. This body played an important role during the first year of implementation of the evaluation policy, contributing to establish the independence of the evaluation office. But it is not fulfilling its role in terms of providing advice to the Evaluation Director, it absorbs significant time from IEO's Directorate and staff for briefings, its reports show no evidence of oversight of the evaluation function and, most important, given that the AEAC reports to the Administrator and it has a role of oversight of the IEO's Director, this institutional design is inconsistent with the independence of IEO (at the end of section 3.3 there are additional considerations about the AEAC).

3.3. Discussion of the key constraints related to the evaluation architecture

Production and Absorptive Capacity Constraints: With the commitment of 100% coverage of country programmes by ICPEs as opposed to partial coverage, IEO's Country Programme Evaluation Section moves away from a two programme cycle consideration to one, shorter in-country missions, and a narrow focus on capturing lessons based on three key questions to inform new country programme strategies. It remains to be seen whether quality can be maintained with the shift in approach and the current level of staff capacity, relative to the overall volume of evaluations. The Country Programme Evaluation Section represented 28% of the annual total budget,⁸ which is a large share of the budget allocation relative to other sections in IEO. Although a number of staff expressed concern in delivering 100% in a timely manner without sacrificing quality,⁹ at this stage, it is still too early to make a reasonable assessment.

It is to be noted that if the production challenge is overcome, and a strong increase in the number of evaluations is produced, then a management absorptive capacity constraint could be generated as it is likely that management may have limitations to deal in a significant, learning inducing way with the process of providing a management response. So the challenge is not only producing more evaluations without sacrificing quality and timeliness but also ensuring that these additional evaluations do not become merely outputs but are also a means to reach development outcomes through the generation and dissemination of evidence-based knowledge and with processes in place that facilitate the use of evaluations as instruments for learning. A good and recent example is the case of the poverty evaluation and UNDP's management request to be allowed to delay its management response in order to have time to engage into a meaningful dialogue conducive to learning.

⁸ UNDP Evaluation Annual Report (2017).

⁹ The number of ICPEs planned in 2018 was 15, whereas the number planned for 2019 is 37.

Quality of decentralized evaluations: A quality assessment of 261 evaluations completed in 2017 highlighted a decline in quality between 2016 and 2017. In particular, the percentage of evaluations with a satisfactory rating fell from 28% to 20%. Indeed, almost half of the individuals interviewed raised the issue of quality of decentralized evaluations, and nearly a quarter mentioned a need to address risk of a conflict of interest in an effort to improve quality. The main organizational challenges for carrying out decentralized evaluations include the low capacity of M&E specialists and regional advisors, limited time devoted to the management of evaluations on the part of M&E specialists, limited resources to conduct decentralized evaluations, uneven access to qualified evaluation expertise, and the risk of the conflict of interest with managers in Country Offices hiring external consultants.¹⁰ However, the issue of conflict of interest would require additional oversight, in line with para. 37 of the Policy, whereas the evaluation capacity problems could require a bridge in the evaluation architecture, linking regional M&E specialists with IEO.

As the variability in the quality of decentralized evaluations is a well-known problem,¹¹ often (but not exclusively) linked to the quality of the evaluators depending on the region, the IEO could further enhance ERC search function by listing the quality assurance scores next to the reports and include a sorting function in the site which orders reports from highest to lowest quality assurance scores, and also an option to sort by consultant and quality assurance scores. This type of public transparency may encourage Country Offices to improve not only the quality of decentralized evaluations but also reinforce oversight without imposing onerous bureaucratic structures. While the ERC is a searchable repository site for different types of UNDP evaluations, it could improve on the user-experience,¹² especially if regional bureaux and country offices want to use the site as a resource to build a quality roster of evaluators.

Concerning the AEAC: This committee is part of the oversight mechanism for IEO. Membership was expanded to include two evaluators, but the majority are auditors. The committee played a valuable role when the policy started to be implemented in safeguarding the independence of the office. But at this stage, the AEAC provides very limited, if any, insights on the evaluation policy, IEO's work plan, or advice on UNDP's decentralized functions and national capacity programming, as revealed by a review of the minutes of the AEAC and interviews.¹³ Additionally, as the AEAC also reports to the Administrator, this structure creates the potential to jeopardize the independence of the Director of IEO. For UNDP's evaluation function the AEAC has morphed from an asset to a liability.

¹⁰ Creating a perverse incentive to make positive evaluations of their own projects.

¹¹ *Peer review of UNDP's Evaluation Office on methodology and knowledge sharing* (15 January 2013); Baastel, "Review of the UNDP Evaluation Policy: Final Report," (8 October 2014).

¹² Not all the information tabs on the ERC are filled out or updated regularly. In fact, one must click quite a few times to find names and contact information of evaluators and then compare separately if the report is of acceptable quality.

¹³ While interviews indicated mixed consensus of those who found the AEAC useful as an oversight function vis-à-vis those who did not, the minutes show that the actual time devoted to substantive evaluations issues relative to audit and other issues is minimal.

4. Conclusions and Recommendations

The preceding analysis of strengths and weaknesses of UNDP's Evaluation Policy and the Evaluation Architecture can be summarized in a set of conclusions that point the way towards recommendations to address the problems identified.

4.1. Conclusions

- a) *A well-crafted and useful policy:* The evaluation policy approved by UNDP's EB in September 2016 provided a framework that allowed for the consolidation of an independent evaluation office. The implementation of the 2016 Evaluation Policy is still at an early stage to assess its effects, and in addition the context for evaluation in UNDP has changed substantially and in a very positive direction, making it even more difficult to attribute changes (or even contribution to changes) to the evaluation policy.
- b) *An evaluation architecture with many strong elements:* The roles and responsibilities of the different elements of the architecture is well described in the policy. It is a broad assessment that IEO has strengthened its role and quality of work during the implementation of the policy. The 2016 UNDP Evaluation Policy, jointly with the Charter of the IEO and the 2019 Evaluation Guidelines, constitute an appropriate framework for an evaluation function that can contribute effectively and efficiently to make UNDP a learning and accountable organization, enhancing its development effectiveness.
- c) *A need for a few adjustments:* Based on the experiences during the implementation of the policy and related to the changed context for the evaluation function, we recommend a few adjustments to be made in the policy and the evaluation architecture.

4.2. Recommendations

An amendment to the 2016 UNDP Evaluation Policy should include a reference to the Charter of the Independent Evaluation Office and to the 2019 Evaluation Guidelines. Furthermore, it would be appropriate for the amendment to introduce the following *changes in sections III-V of the Evaluation Policy*

Principles (section III)

1. The principles of the Evaluation Policy should include an explicit reference to the Agenda 2030, gender equality, diversity, inclusion, human rights and private sector.

Evaluation procedures and quality assurance (section IV)

2. The planning process should involve consultation with stakeholders and in all phases of the evaluation process it is important that evaluators engage with stakeholders and ensure not only the national context is considered but also that

the purpose, relevance and messages of the evaluations are communicated clearly and using a language that does not create unnecessary tensions.

3. The decision on what to evaluate should be made with an explicit statement of the purpose and potential use of the evaluations for strategic decision-making.
4. A technical reporting line of Regional M&E specialists to the IEO Director on evaluation issues would contribute to enhance the quality of decentralized evaluations. Additional funding from evaluation partnerships may be instrumental in developing arrangements to strengthen the oversight and support to decentralized evaluations and to make more useful the ERC, with a better use of QA scores. The evaluation responsibilities of the M&E Regional specialists should be enhanced.
5. The use of different and new types of evaluations and data collection methods should be encouraged, including whenever appropriate a complexity and systems approach and paying attention to innovation and scaling-up.
6. After producing evaluations, efforts should be made to elaborate messages derived from the evaluations, including syntheses, showing trends/patterns based on granular data, that may be of interest to different audiences.
7. The requirement in the Evaluation Policy that a Management Response should be prepared for *all* evaluations and in a fix period of time could be changed so as to alleviate the stress on management's absorptive capacity: in the case of ICPEs the new Country Program could be considered as an option of a Management Response, whereas the EB could allow extensions in the submission of Management Responses.
8. As country program evaluations are no longer mandatory, whereas ICPEs are in the process of arriving at 100% coverage, there could be some flexibility in the overall 1% of allocation of funds for evaluation, by introducing a link between the 0.8% for non-IEO evaluations and the evolution of UNDP's portfolio of activities and funds. This would be facilitated if UNDP introduces a budget line to accurately capture funds allocated to evaluation. The ambivalence regarding the funding level should be eliminated by deleting the last part of the sentence in paragraph 26 "subject to availability".

UNDP Evaluation Architecture (Section IV)

9. Given the structural problem with respect to independence posed by the existence of an AEAC that reports to the Administrator, potentially compromising the independence of the IEO Director, the AEAC should no longer be part of UNDP's evaluation architecture.
10. An independent and external review of the evaluation function should be conducted every four years by an external team reporting to the Board.

5. Annex 1

List of Persons Interviewed

Office	Name	Title
SM-OA	Achim Steiner	Administrator of UNDP
ROK	Tae-yul Cho	President of the Executive Board
SM-OA	Michele Candotti	Chief of Staff
SM-OA	Joseph D Cruz	Senior Advisor to Administrator
SM-BMS	Susan McDade	Assistant Administrator
SM-RBA	Ahunna Eziakonwa	Assistant Administrator
SM-BERA	Ulrika Modeer	Assistant Administrator
SM-RBAS	Mourad Wahba	Assistant Administrator
SM-RBAP	Haoliang Xu	Assistant Administrator
SM-BPPS	Mar Dieye	Assistant Administrator
SM-RBEC	Mirjana Spoljaric Egger	Assistant Administrator
SM-OA	Theresa Panuccio	Senior Advisor
SM-BERA	Gulden Turkoz-Cosslett	Deputy Assistant Administrator
SM-RBLAC	Lenni Montiel	Deputy Regional Director
SM-RBAS	Susanne Dam Hansen	Strategic Advisor
SM-BMS, OFRM	Darshak Shah	Director
SM-RBAP	Faiza Effendi	Strategic Planning Advisor
RBA	Mamadou N'Daw	RBM and Evaluation Advisor
RBEC	Tahmina Anvarova	Regional Specialist
RBEC	Ekaterina Paniklova	Senior Programme Coordinator
IEO	Indran Naidoo	Director
IEO	Arild Hauge	Deputy Director
IEO	Heather Bryant	Chief (Capacity dev & QA)
IEO	Alan Fox	Chief (Corporate evaluations)
IEO	Fumika Ouchi	Chief (of ICPEs)
IEO	Richard Jones	Evaluation Advisor/ QA
IEO	Deqa Musa	Evaluation Specialist
IEO	Ximena Rios	Chief (Operations)
SM-BPPS	Margaret Thomas	Chief
SM-BPPS	Kristina Leuchowius	Policy Specialist
SM-BPPS	Adriana Dinu	Deputy Director BPPS
RC Samoa	Simona Marinescu	RC/former Chief Development Impact Group BPPS
BPPS	Nancy Bennett	M&E/Results Management Advisor
UNDP Audit	Helge Osttveiten	Director
OIOS IED	Yee Woo Guo (Eddie)	Director
UNV	Martin Hart Hansen	Chief, Strategic Planning Advisor
UNV	Dominic Allen	Chief, Partnerships
UNV	Lauren Phillips	Partnerships
UNCDF	Andrew Fyfe	Head of Evaluation
UNCDF	Xavier Minchon	Deputy Executive Secretary

Fiji	Peter Thomson	former Ambassador
US	Charles Chang	Foreign Service Officer
Denmark	Ib Petersen	former Ambassador
Baastel	David Todd	Consultant
UK	Emily Braid	Senior Policy Advisor
AEAC	Sheila Fraser	Chair
AEAC	Mallika Samaranayake	Evaluation Committee member
AEAC	Ryokichi Hirono	Evaluation Committee member
UNFPA	Marco Segone	Director of Evaluation
UNICEF	Riccardo Polastro	Regional Evaluation Adviser
IMF	Ruben Lamdany	Deputy Director Evaluation

*A focus group was carried out with IEO staff P5 and below including General Service.

Breakdown of Interviewee by professional type and gender	
Total number of individuals interviewed ¹⁴	51
Senior management (Administrator, ASG, D2, D1, and P6 levels)	24
UN system Professional level staff (P5 to P3)	17
Other non-staff (Permanent Representatives or Committee Members)	10
Male to Female ratio	28 males 23 females

¹⁴ Excluding 1 focus group (with a total of 13 IEO staff).

6. Annex 2

Methods

A qualitative approach was employed to address the evaluation policy review questions. First a desk review was carried out to identify potential issues and topics for the development of the interview guide. The selection process for persons interviewed was based on the level within the organization, degree of insight on the 2014 and 2016 policy review processes, key stakeholders (including funds and programmes administered by UNDP, UNDP staff outside of IEO, UN staff, members of the Executive Board (current and former) and regional bureaux coverage). Given the short duration of this review period, systematic coverage of UNDP country offices was limited.

Semi-structured interviews were then conducted by the Review Team between 4 to 28 February 2019. The semi-structured interviews were guided by the interview templates (below) and limited to a 1-hour duration during which follow up questions were permitted. The interviews were transcribed and reviewed amongst the Review Team to cross-check content accuracy and used for team debriefing.

A preliminary map of thematic issues (see: Figure 4) was built based on a batch of interview transcripts. The red circles denote the emergent thematic issues and all other colors related sub-themes. All the interview transcripts were subsequently analyzed using open and axial coding. Codes were assigned units of meaning (i.e., portions of text varying size such as words, sentences, and phrases) in the transcripts. Open coding (fracturing of data and grouping/categorizing) was performed, followed by iterative axial coding (rearranging the data in new ways). This coding was completed manually twice and then cross-checked against a QDR software using a cluster analysis of themes (see: Figure 5). Multiple themes were identified in the transcripts and grouped under broader themes and discussions. The final thematic clusters are depicted in Figure 6, which form part of the basis of analysis in the review.

Figure 4: Emergent themes, phase I

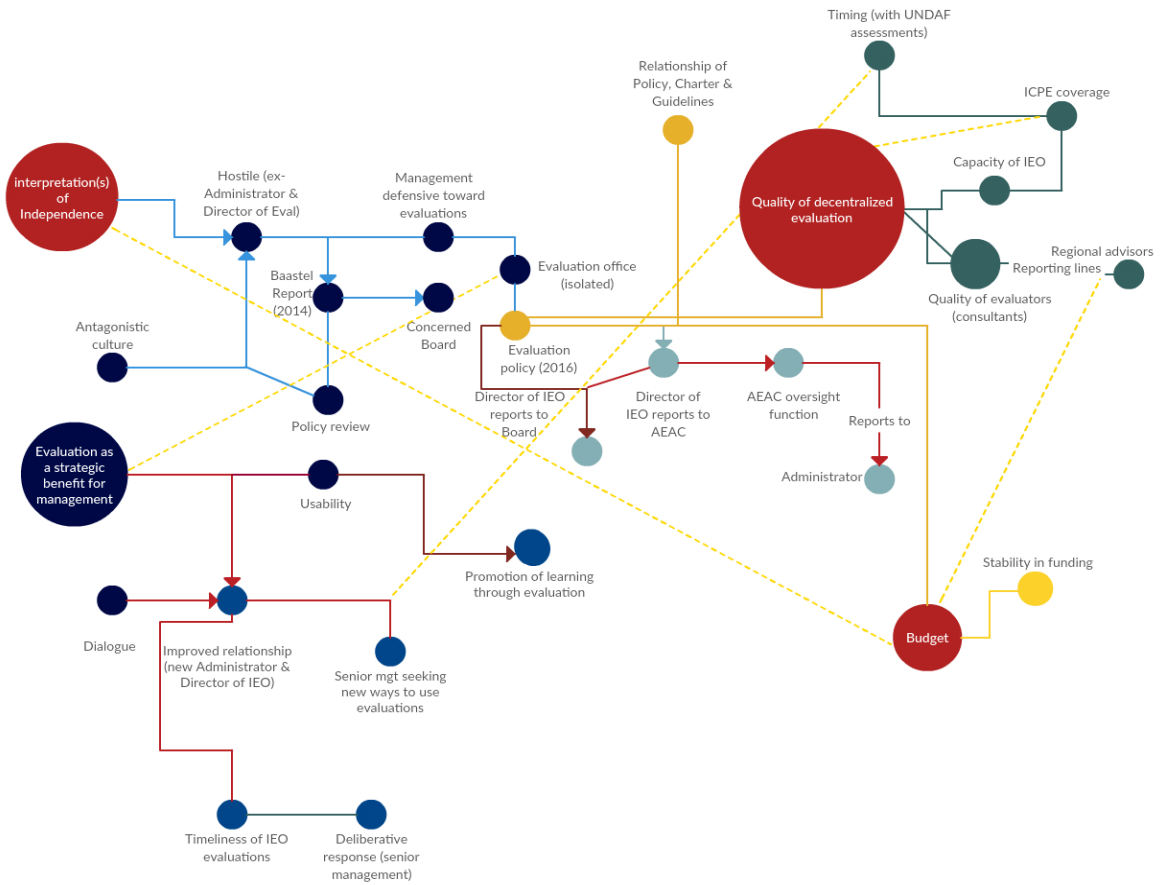


Figure 5: Coding co-occurrences and similarity index, phase II

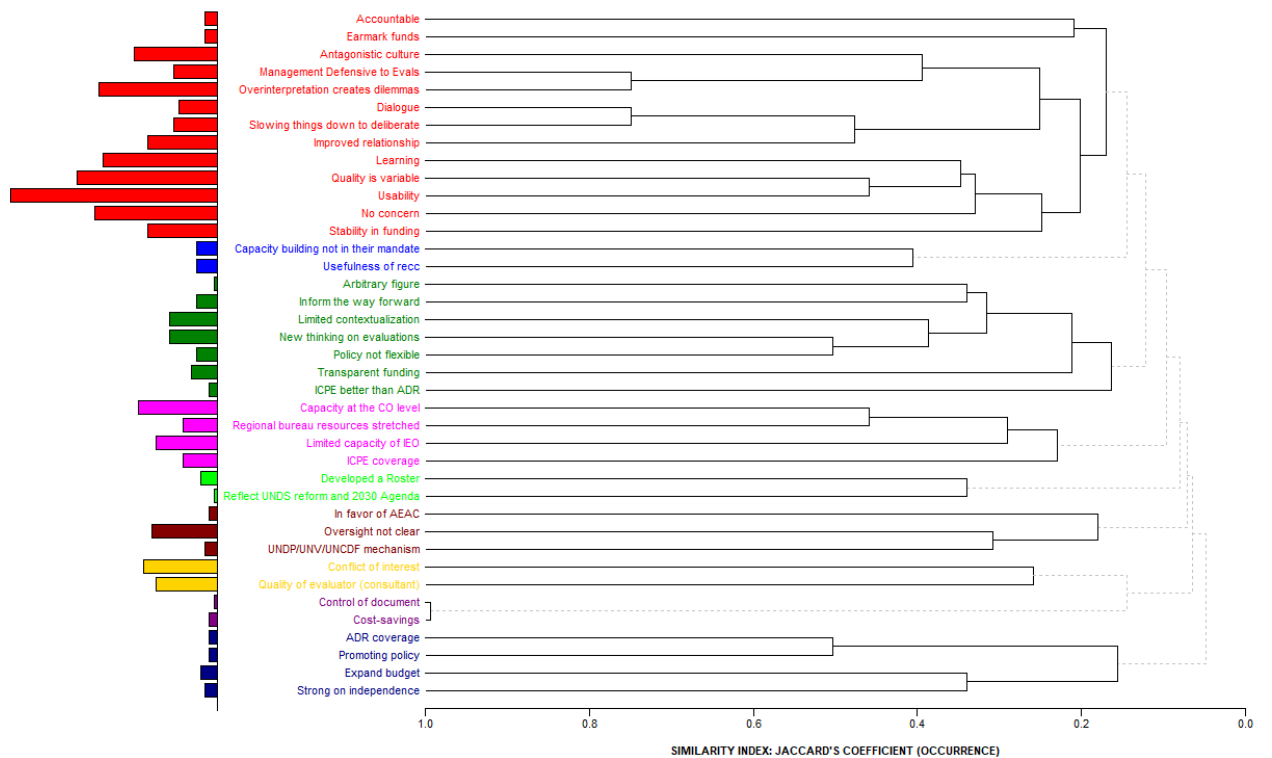
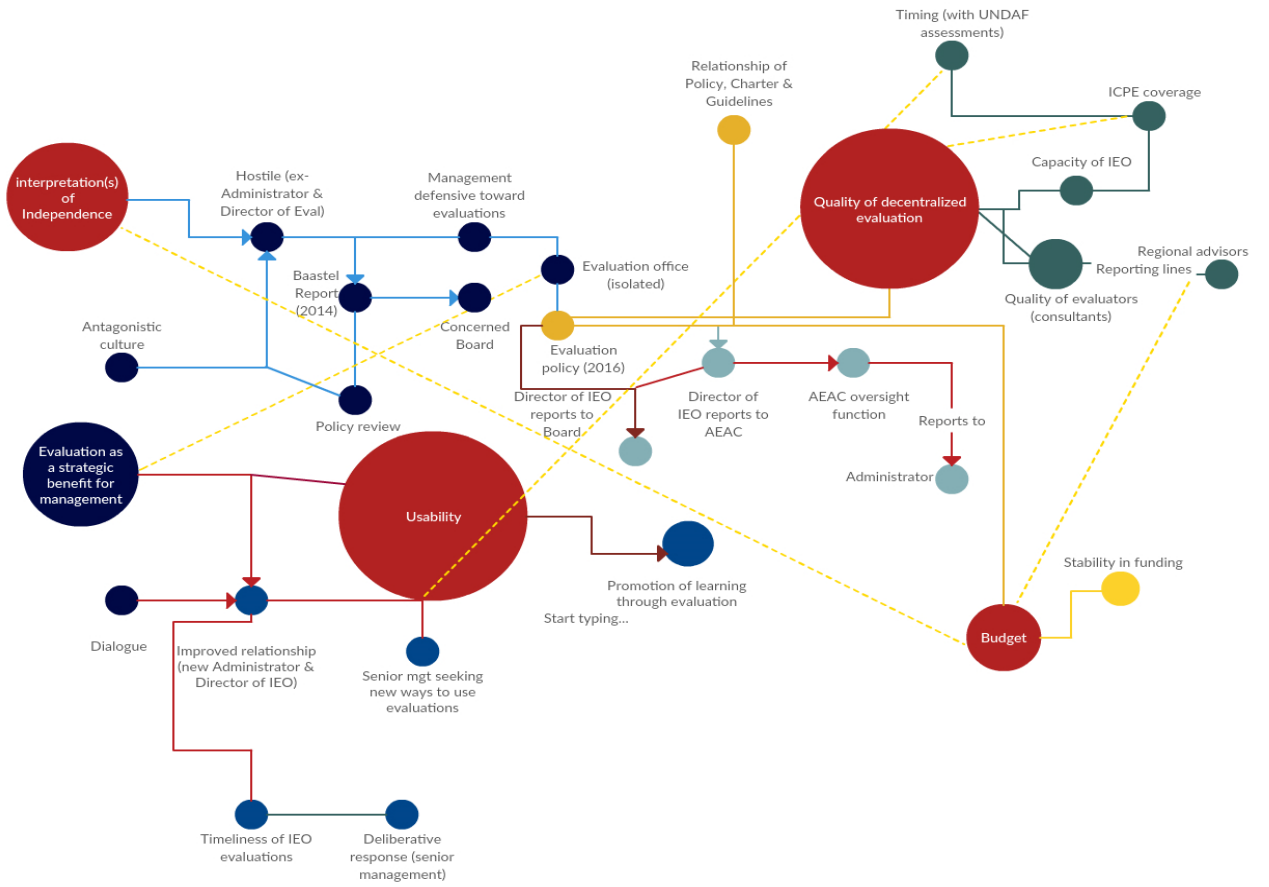


Figure 6: Emergent themes, phase III (final)



Questionnaire for Semi-Structured Interviews [General template]

Have you seen UNDP's 2016 Evaluation Policy?

With respect to independent evaluations at UNDP, do you perceive that since 2016 the following aspects has been improved or weakened:

1) credibility? 2) independence? 3)use? Please explain

With respect to decentralized evaluation, is there, in your view, adequate institutional capacity at UNDP at country, regional and HQ levels to manage and or conduct these evaluations and assured their quality?

What could be done to strengthen the capacity further?

Do you think that the budget allocated for evaluations is enough? Do you think evaluation plans are fully financed?

Regarding independent and decentralized evaluations

In your view, are the recommendations from the independent evaluations being followed up by management and the board?

What can be done to strengthen evaluation follow-up?

Were there improvements in UNDP's systems and practices that can be attributed to independent evaluations done according to the 2016 evaluation policy?

Please refer to evidence and supporting documentation.

To what extent has the 2016 UNDP evaluation policy contributed to guiding decisions and actions when conducting evaluations?

What do you consider as the strengths/weaknesses of the policy?

What is your overall assessment of the evaluation policy?

What do you think should be improved in the policy?

Regarding independent &/or decentralized evaluations

Regarding evaluation capacity development

Questionnaire for Semi-Structured Interviews - UNCDF

Have you seen UNDP's 2016 Evaluation Policy?

With respect to independent evaluations at UNCDF, do you perceive that since 2016 the following aspects has been improved or weakened:

credibility? 2) independence? 3)use? Please explain

With respect to decentralized evaluation, is there, in your view, adequate institutional capacity at UNCDF at country, regional and HQ levels to manage and or conduct these evaluations and assured their quality?

What could be done to strengthen the capacity further?

Do you think that the budget allocated for UNCDF evaluations is enough? Do you think UNCDF evaluation plans are fully financed?

Regarding independent and decentralized evaluations

In your view, are the recommendations from the independent evaluations being followed up by management and the board?

What can be done to strengthen evaluation follow-up?

Were there improvements in UNCDF's systems and practices that can be attributed to independent evaluations done according to the 2016 evaluation policy?

Please refer to evidence and supporting documentation.

To what extent has the 2016 UNDP evaluation policy contributed to guiding decisions and actions when conducting evaluations?

What do you consider as the strengths/weaknesses of the policy?

What is your overall assessment of the evaluation policy?

What do you think should be improved in the policy?

Regarding independent &/or decentralized evaluations

Regarding evaluation capacity development

Questionnaire for Semi-Structured Interviews - UNV

Have you seen UNDP's 2016 Evaluation Policy?

With respect to independent evaluations at UNV, do you perceive that since 2016 the following aspects has been improved or weakened:

credibility? 2) independence? 3)use? Please explain

With respect to decentralized evaluation, is there, in your view, adequate institutional capacity at UNV at country, regional and HQ levels to manage and or conduct these evaluations and assured their quality?

What could be done to strengthen the capacity further?

Do you think that the budget allocated for UNV evaluations is enough? Do you think UNV evaluation plans are fully financed?

Regarding independent and decentralized evaluations

In your view, are the recommendations from the independent evaluations being followed up by management and the board?

What can be done to strengthen evaluation follow-up?

Were there improvements in UNV's systems and practices that can be attributed to independent evaluations done according to the 2016 evaluation policy?

Please refer to evidence and supporting documentation.

To what extent has the 2016 UNDP evaluation policy contributed to guiding decisions and actions when conducting evaluations?

What do you consider as the strengths/weaknesses of the policy?

What is your overall assessment of the evaluation policy?

What do you think should be improved in the policy?

Regarding independent &/or decentralized evaluations

Regarding evaluation capacity development

7. Annex 3

Terms of Reference

Terms of Reference

Independent Review of the UNDP Evaluation Policy

(16 December 2018)

Background

This Terms of Reference is developed for the purpose of carrying out a review of the UNDP Evaluation Policy. The first UNDP Evaluation Policy was approved by the UNDP Executive Board in 2006, and subsequently revised in 2011. Following an independent review of the Policy in 2014, the next and current iteration of the UNDP Evaluation Policy was approved by the Board in September 2016. In its decision approving the 2016 Policy, the Board requested that the Independent Evaluation Office of UNDP commission a review of the policy in 2019.

The UNDP evaluation policy sets out the purpose and basic principles of evaluation for UNDP and its associated funds and programmes.¹⁵ The Policy defines the institutional architecture and clarifies the roles and responsibilities within the UNDP institutional framework, providing a policy foundation for safeguarding the independence of evaluations. The Policy covers all evaluations conducted by the Independent Evaluation Office of UNDP, as well as those commissioned by programme and policy units, and by the associated funds and programmes. The policy harmonizes the UNDP oversight functions under an expanded Audit and Evaluation Advisory Committee (AEAC). The Evaluation Policy establishes financial benchmarks for the UNDP evaluation function, stipulating that 1% of core and non-core resources are to be set aside for the evaluation function annually, within which 0.2% is for the work of the Independent Evaluation Office of UNDP.

Purpose, Scope and Objectives

The review will cover the period from September 2016 to January 2019, taking into consideration contextual and organizational changes since the approval of the 2016 Evaluation Policy. The review will encompass a select set of key aspects of the evaluation policy. This review will:

- assess the progress made in implementing the revised evaluation policy of 2016, noting strengths and weaknesses
- review the evaluation architecture
- identify any constraints inhibiting the effective implementation of the policy and areas that may require policy change.

The findings and recommendations will be presented to the UNDP Executive Board and UNDP management during the annual session of the Executive Board in June 2019.

Review Questions

The following questions are established for the team to address through the review.

a) Implementation of the Evaluation Policy

¹⁵ UN Volunteers (UNV), and the UN Capital Development Fund (UNCDF)

Has the 2016 policy influenced the systems and practices of UNDP, as well as UNV and UNCDF?

- To what extent is the evaluation policy known?
- Is there evidence of improvement in the independence, credibility and use of evaluation at UNDP and the associated funds and programmes as a result of the revised policy?
- Is there adequate institutional capacity to meet the evaluation policy requirements at UNDP and the associated funds and programmes, at country, regional and HQ levels?
- Have the financial benchmarks set in the policy been met? Are the evaluation plans for UNDP and the associated funds fully costed?
- Are UNDP and the associated funds and programmes taking action in response to evaluation recommendations?

b) Evaluation Policy content

Based on the review analysis, taking into account answers to the above questions, and changes in the context in which UNDP and the associated funds and programmes operate, are there clarifications and improvements that should be made to the existing policy text under its respective headings: Principles, Procedures, Architecture?

Approach and Methodology

The review will not be a full-fledged evaluation; however the review team is expected to take into account UNEG Norms and Standards for Evaluation. The Review should include:

- Desk Review of selected reports
- Individual and group interviews (in person and phone)
- One-week visit in New York in February 2019 for interviews with key informants and stakeholders.

Expected Deliverables

- Draft and final reports (no more than 30 pages excluding annexes), submitted by the end of April 2019. The main report will cover the methodology, main findings, conclusions and actionable recommendations.
- Summary paper (up to 8000 words) submitted by the end of April 2019, for distribution to the Executive Board.
- Participation in an informal presentation of the review results to the Executive Board in May/ June 2019

Team Composition and Responsibilities

The Review Team includes three individuals: Team Leader, Senior Advisor, and Senior Consultant.

The Team Leader, Osvaldo Feinstein, is a member of the UNDP/IEO Evaluation Advisory Panel. Osvaldo (10 days) will coordinate the effort, finalize the approach and methods, lead the factfinding mission to UNDP HQ in New York, and lead the review findings presentation to the Executive Board.

The Senior Advisor is Per Øyvind Bastøe, Director of Evaluation at the Norwegian Agency for Development Cooperation (NORAD, Ministry of Foreign Affairs) and Chairman of the OECD/DAC

Evalnet. Per (10 days) will support the effort in an unpaid capacity, participate in the factfinding mission, help to develop the draft review report and participate in the presentation to the Executive Board.

The Senior Consultant is Patty Chang, Adjunct Associate Professor at NYU Center for Global Affairs. Patty (20 days), will contribute to the design of the review, conduct background documentation analysis, participate in the factfinding mission, and support the Team Leader in drafting and then presenting the review.

Implementation Arrangements

The Independent Evaluation Office will support the review, including assisting the review team by facilitating interviews, tracking down background documentation; and arranging for the publication and dissemination of the review report. The IEO Chief of Corporate Evaluation, Alan Fox, will take responsibility for internal IEO management aspects, supported by IEO professional and operational staff for research, contract management and travel facilitation.

UNDP management will identify a focal point to liaise with the team during the review, in particular as relates to decentralised evaluation aspects. Likewise, a focal point for this exercise will be identified at UNCDF and UNV. Management of UNDP, UNV and UNCDF will have an opportunity to comment on the draft final report prior to completion. Management will ensure that the review team receives needed support for data collection and that factual comments on the draft review report are received in a timely manner.

Timeframe

The review will commence 15 December 2018, with a draft report submitted by 5 April 2019. Following review and revision, the final report is due by 26 April 2019.

A visit to New York for the Evaluators is scheduled for the 1st week in February 2019, for briefings and interviews with UNDP senior management and other HQ stakeholders. A second visit is envisaged, in May/June for the team to participate in a UNDP Executive Board informal where this matter will be taken up.