

8 January, 2007

**DAC/UNEG Joint Task Force on Professional Peer Reviews of Evaluation Functions in  
Multilateral Organizations**

**Framework for Professional Peer Reviews**

**Contents**

1. Summary.....	1
2. Background and rationale for the professional peer reviews.....	3
3. Approach and methodology: Recommendations and options .....	4
4. Roles and responsibilities in the reviews.....	10
5. Phase-by-phase organization of a professional peer review process.....	12
6. Time and resources required.....	15
7. Planned monitoring and assessment of experience.....	15

**Summary**

1. This framework for professional peer reviews of evaluation in multilateral organizations is one contribution to the efforts in the international community to strengthen a shared agenda for improving performance in development and cooperation. Since 2004, the community of evaluators in international cooperation agencies have themselves developed and tested (through pilot peer reviews in UNDP and UNICEF) a new approach to assessing and enhancing multilateral agencies' own evaluation capacity and performance. Now the DAC Evaluation Network and the UN Evaluation Group have joined forces to establish an internationally recognised form of peer review of the evaluation function. The joint task force will focus on developing a framework for professional peer reviews, based on previous experiences and internationally recognized standards.

2. This approach has several purposes: building greater knowledge, confidence and use of evaluation systems by management, governing bodies and others; providing a suitable way of "evaluating the evaluators"; sharing good practice, experience and mutual learning. The primary intended audience for the results of these professional peer reviews is one of decision-makers and other users of evaluation – including where appropriate the intended beneficiaries in member countries.

3. On the basis of substantial analysis and lessons from the pilot experiences, this framework sets out both specific recommendations and some options for the "**Approach and Methodology**" for these professional peer reviews, beginning with an explanation of the nature, strengths and limits of a peer review. Because the peer reviews are intended to assess the evaluation function against accepted international standards in a wide range of organizations, the framework recommended here has a blend of standard and flexible elements. It sets out the specific focus of the peer reviews of evaluation.

4. The framework contains clear guidelines on a consistent, broad normative framework using applicable norms and standards, and also provides some tested, step-by-step options and practical suggestions for carrying out the review. At the same time, it is intended to allow for necessary flexibility, for example to adapt for de-centralised and centralized evaluation systems or to feature special issues or emphases for particular organizations. It is worth noting that while

the framework has been primarily designed for reviews of multilateral development agencies, it should prove equally useful (with limited adaptation) for any evaluation function such as those of international financial institutions, individual bilateral donor agencies, or other multilateral agencies, for example those concerned with humanitarian relief or human rights issues.

5. The framework points out there are a number of issues that are likely to require decision as the plan is finalized for each individual review exercise. It is also clearly recognized that work may need to be done to adapt the model much further in some cases - for example, to very small or highly specialized organizations, and/or those with extremely limited existing evaluation capacities. The framework also includes an explicit treatment of the risks that may arise in these reviews, and ways of managing them.

6. The Framework for Assessment is outlined, including a core assessment question. As reflected in this question, the approach and methodology hinges on using recognized normative frameworks and clustering the treatment of the many issues under three crucial criteria for evaluation, specified as:

1. "The independence of evaluations and evaluation systems"- Noting that this attribute is never absolute, and needs careful analysis in different contexts;
2. "The credibility of evaluations". This includes assessment of whether and how the organization's approach to evaluation fosters partnership and helps build ownership and capacity in member countries where appropriate; and
3. "The utility of evaluations" – Recognizing that this is only partly under the control of evaluators, and is also critically a function of the interest of managers, and member countries through their participation on governing bodies, in commissioning, receiving and using evaluations.

7. Depending on their applicability to the organization being reviewed, one of several sets of accepted and consistent standards should be applied to provide a normative framework for the assessment under these three broad headings. To avoid future professional peer reviews of UN organizations having to repeat the necessary sorting and categorization work needlessly, such review teams are referred to the documentation of the UNDP and UNICEF peer reviews to be able to extract readily the relevant materials, and similar work was done for the UN Norms and Standards and the DAC Principles. These peer reviews may be one of the most important continuing general tests of the usefulness of the prevailing norms and standards for evaluation, and they should thus be prepared to suggest ways to strengthen the norms on the basis of these experiences.

8. The experience of the pilot reviews suggests that it may normally be appropriate to apply a combination of indirect (evaluation system) and direct (sample evaluation product) assessments at different levels. Although there are not yet single, internationally accepted standards for assessing the quality of evaluation products, a number of useful models are already being tested and applied at different levels.

9. Given the central importance of the Normative Framework to the whole peer review process and its results, experience confirms the importance of the specific framework being agreed between the peer review team and organization at an early stage.

10. The framework outlines the key roles and responsibilities of different parties in the professional peer reviews, beginning with the peer Panel, including appropriate criteria for selection of its members, and the main expectations for its tasks. Similarly, it outlines the major roles and responsibilities that will fall on the organization under review, which will normally be carried mainly, but not exclusively, by its evaluation unit.

11. To facilitate the work of those launching future reviews, the main lines of the approach, and some important practical lessons, are illustrated with typical steps and sequences, in a section providing a “phase-by-phase organization of a peer review process”. The Framework document concludes with brief discussions of the time and resources that are likely to be required for a professional peer review of evaluation, and of the planned monitoring and assessment of experience in these reviews.

## **1. Background and rationale for the professional peer reviews**

### *Background*

12. In the field of international cooperation, the international community has come together around a shared agenda for improving performance in development and cooperation. As part of this drive, the need for strong monitoring and evaluation has been a key concern. It is closely linked to setting clearer objectives, improving working practices, and applying the lessons of experience - all with member countries and their people more firmly in the “driver’s seat”, where they must be if durable progress is to be achieved.

13. All the key actors in international cooperation are now part of this drive for evidence of effectiveness and improvement, including the multilateral development agencies that channel a large share of total assistance flows to developing, transitional and other member countries. The member countries of these agencies need credible evidence of effectiveness and benchmarks for improvement, as do their partners on the ground, as do their own senior managers and operational staff. At the same time, assessing the effectiveness of these organisations can pose particular challenges because of their wide membership, international governance and sometimes the range and types of their activities.

14. After the experience in recent years of several major external evaluations of multilateral agencies – mostly organised by consortia of funding governments – two communities of evaluators in development agencies<sup>1</sup> have since 2004 themselves developed and tested a new approach to assessing and enhancing multilateral agencies’ own evaluation capacity and performance. It is worth noting that while the framework has been primarily designed for reviews of multilateral development agencies, it should prove equally useful (with limited adaptation) for any evaluation function such as those of international financial institutions, individual bilateral donor agencies, other multilateral agencies, for example those concerned with humanitarian, relief or human rights issues, as well as global programs, foundations and global non-governmental organizations.

### *Rationale*

15. This approach has several purposes:

---

<sup>1</sup> The OECD/DAC Network on Development Evaluation, and the UN Evaluation Group (UNEG).

1. Building greater knowledge and use of evaluation systems by management, governing bodies and others, leading to: better understanding of the current quality and needed improvements in evaluation, better integration of evaluation priorities and findings into performance management, improved evaluation policy and practice at all levels, and stronger planning and budgeting for evaluation;
2. Within an organization, providing a way of “evaluating the evaluators” through a professional peer assessment against recognized international standards, thus respecting the necessary degree of independence of the evaluation function from direct assessment by management;
3. Sharing good practice, experience and mutual learning, building greater internal capacity and external confidence in multilateral evaluation systems and thus ultimately reducing demands for special outside assessments of performance.

16. On the invitation of the Evaluation Offices of UNDP and UNICEF, pilot peer reviews of evaluation in those two organizations were carried out and published in 2005-2006, and the experiences of both exercises were documented and assessed by the participants.<sup>2</sup> Both the pilot reviews were found to have been extremely valuable, in distinct ways, for the two organizations and panels concerned.

17. The DAC Evaluation Network and the UN Evaluation Group have now joined forces to establish an internationally recognised form of professional peer review of the evaluation function. The joint task force has focused on developing this framework for professional peer reviews, based on previous experiences and internationally recognized standards. It will promote peer reviews, starting with interested UN organizations. The joint task force sees its work as catalytic in ensuring that peer reviews will incorporate perspectives from a broad range of professional expertise in (inter)national organizations, including perspectives of member countries, and non-governmental and private sector expertise.

18. *Audience.* In view of the purposes and experience described above, the primary intended audience for the results of these professional peer reviews is one of decision-makers and other users of evaluation – including where applicable the intended beneficiaries in member countries - although there is much to be gained from the exchange between evaluation specialists. Given this primary audience, it is especially important to avoid engaging or appearing to engage in special pleading or professional courtesy among professional evaluators, and to ensure that the review embodies objective and demanding standards for advancing organizational performance and accountability.

## **2. Approach and methodology: Recommendations and options**

### **The overall approach**

*The nature, strengths and limits of a professional peer review.*

19. Peer review can be described as the systematic examination and assessment of the performance of an organization by its peers, with the ultimate goal of helping the reviewed organization improve its policy making, adopt best practices, and comply with established standards and principles. The examination is conducted on a non-adversarial basis, and it relies

---

<sup>2</sup> See references 2 and 3

heavily on mutual trust among the organizations involved in the review, as well as their shared confidence in the process.

20. It is not intended to serve as a procedure for resolving differences and peer review never implies a punitive decision or sanctions; it generally goes beyond fact-finding to include an assessment of the performance, and is characterized by dialogue and interactive investigation. Peer pressure does not take the form of legally binding acts, as sanctions or other enforcement mechanisms. Instead, it is a means of peer persuasion which can become an important driving force to stimulate organizations to change, achieve goals and meet standards.<sup>3</sup>

21. All professional peer reviews contain the following structural elements, and the following sections of this framework outline recommendations and options for applying them to these reviews of the evaluation function in multilateral organizations:

- A basis for proceeding: including the necessary agreement for cooperation, and a clear question for assessment;
- A normative framework: an agreed set of principles, standards and criteria against which performance is to be reviewed;
- Designated actors and roles in carrying out the peer review; and
- A set of procedures for planning, assembling and testing the base of evidence and findings, leading to the final result of the peer review – an exchange around the conclusions drawn primarily from the diverse professional experience and judgments of the peer panelists themselves, deliberating as a group.

22. The device of professional peer assessment, already well tested in the development field, has been agreed to bring a number of special strengths to this undertaking, as illustrated in international processes such as those of the OECD, WTO and now NEPAD. First, it starts with a shared appreciation of the distinctive challenges of work and evaluation in international cooperation, and the fact that all concerned are constantly striving to improve. Second, it can adapt and apply the most pertinent professional principles, norms and standards in coming to an assessment. Third, drawing on experienced professional peers from other institutions (participating as individuals) will maximize the opportunities for sharing relevant experience and lessons. Finally, the resulting assessment should carry particular weight, both internally and externally, for the independence and professional credibility of its results. All these aspirations highlight the importance of the quality of the panel and its support, as well as the openness and cooperation of the organization being reviewed.

23. To avoid unrealistic expectations, it is important to be clear that a professional peer review of evaluation in an organization is not in itself an assessment of the effectiveness of that organization. However, it can and should contribute to the basis for assessing the effectiveness of the organization, by testing the capacity and quality of the organization's own evaluations of effectiveness, and thus the confidence that can be placed in them. Further, while a peer review of

---

<sup>3</sup> The basic parameters of peer review outlined here are taken from "Peer review: a tool for co-operation and change: An analysis of an OECD working method," by Fabrizio Pagani. Organisation for Economic Co-operation and Development, Paris, 2002. They have been adapted to the special character of reviews under this framework.

the evaluation function is focused on one important part of an organization's performance management systems, at the same time it is also likely to shed light on other parts, such as the framework for managing for results, etc. Maintaining this specialised focus on evaluation will be important for the reviews, while also taking into account how it actually works with these inter-related systems in the organization.

24. It should also be stressed that these peer reviews:

- Are not themselves full-fledged evaluations – they are less comprehensive and in-depth assessments but, as outlined below, they should adhere to a rigorous methodology – applying the key principles of evaluation while taking full advantage of the particular benefits of a peer mechanism. It is explicit that the final conclusions will clearly be a judgment by the Panel concerned.
- Must themselves reflect the accepted standards of good practice in development cooperation by seriously engaging developing country partners in the process of the review, in the learning acquired, and in the use of the results.
- Are designed to be targeted and lean – without sacrificing the rigour required for their basic validity and credibility – to avoid making unreasonable demands of time, expense and additional workload on both Panel members and the organization being reviewed.

25. Because the professional peer reviews are intended to assess the evaluation function against accepted international standards in a wide range of organizations, the framework recommended here has a blend of standard and flexible elements. Thus it contains clear guidelines on a consistent broad normative framework using applicable norms and standards, and also provides some tested, step-by-step options and practical suggestions for carrying out the review. At the same time, it is intended to allow for necessary flexibility for example, to adapt for de-centralised and centralized evaluation systems or to feature special issues or emphases for particular organizations. Thus there are a number of issues that are likely to require decision as the plan is finalized for each individual review exercise. It is also clearly recognized that work may need to be done to adapt the model much further in some cases - for example, to very small or highly specialized organizations, and/or those with extremely limited existing evaluation capacities.

26. Finally, because the approach is new and still being refined it is not intended, to be overly prescriptive. Experimentation is to be expected, and the documentation and sharing of experience and lessons is encouraged, with a view to possible further formalization (e.g. rotational coverage vs. volunteering) on the basis of the experience of a substantial number of further reviews.

#### *Risks and their management*

27. Among the possible risks that may be encountered in carrying out these professional peer reviews, four appear to be most prominent:

- i. The review might become too heavy to be justified – e.g. very burdensome on the organization being reviewed and/or the panel, or too expensive or extended;
- ii. At the other extreme, it might become too light to be credible – e.g. too cursory or superficial a treatment, or having insufficient engagement of the panel members to warrant the claim of a properly informed peer assessment;

- iii. It might become, or be perceived as, too “cosy” an assessment among professional evaluation peers, lobbying for this professional function or exercising too much “professional courtesy” in its assessments; or
- iv. Because it requires access to sometimes-sensitive information and leads to important assessments of the quality of the work of organizations and teams, each review will encounter and have to manage different views on process and substance, and could become bogged down in disputes at various levels.
- v. It might be perceived as a donor-driven exercise to justify decreased voluntary contributions to the organization concerned if the conclusions of the review are negative;
- vi. At the other extreme it might be perceived as an exercise conducted to provide justification for higher voluntary contribution if the review is very positive.

28. The first three sets of risks identified above have each been encountered and tested in the pilot reviews, and the framework outlined in this paper has been designed to include features to mitigate each of them to the extent possible. At the same time, it will be important for all parties embarking on such a review to register and discuss these risks explicitly at an early stage, and build any necessary risk-mitigation measure into the specific review plan.

29. The fourth risk – that of serious disputes which might come to threaten the integrity or completion of the review – also needs to be anticipated and specifically managed. It should be stressed that differences of approach and interpretation are intrinsic to any such exercise, and that the open “dialogue and interactive investigation” in the process are a major part of its value. Nonetheless, even in very well institutionalized peer review systems, intractable conflicts can arise among participants at various levels. In such cases – always recognizing that the process is a non-coercive one – the normal recourse is to call on the body sponsoring the review system to help resolve or manage the conflict. However, since that body in this case does not have any legal status, it would be preferable to invite the governing body of the organization concerned to initiate a process of mediation. This body may seek advice or support from the co-chairs of the Joint UNEG-DAC Task Force and mediation efforts could include appointing one or more arbitrators to advise the governing body on how to proceed. It is recommended that this contingency should be specifically recognized and provided for in the agreed peer review plan between the panel and the host organization.

30. The last two risks will need to be confronted in the mixture of panel members, where restraint should be applied in appointing members from donors that have a high financial stake in the organization or donors that are in the process of reconsidering the level of their voluntary contributions.

*Focus of the peer reviews of evaluation.*

31. The professional peer review would focus on the evaluation function within the organization, take its central evaluation unit as a starting point and working towards understanding how evaluation operates at the various levels, in order to review the quality of the function in light of the objectives of the organization and the appropriate international standards. Specific aims would be:

1. To assess structural aspects of how the evaluation function operates in the organization. Depending on the organization reviewed, this should look at the evaluation function in

- the perspective of how the organization can better report on and increase its effectiveness in reaching its goal.
2. The peer review should look at policy issues, starting with the evaluation policy of the organization and other related policies, as well as issues of planning, budgeting and coverage.
  3. It should address various relevant levels in the organization, such as the governance level (board, council), central management functions and the central evaluation unit, decentralized evaluations both in headquarters and in field offices, and outsourced evaluations in interventions.
32. It should look at existing capacity and need for capacity strengthening, use of relevant and appropriate evaluation methodologies, and the quality of the evaluations undertaken, in preparation, implementation and reporting. Another key focus should generally be to look at collaboration in evaluation with local partners and stakeholders, as well as harmonization and coordination with other external partners.

### **The Framework for Assessment**

*A core assessment question.*

33. The professional peer reviews of evaluation are to be centred on a “core assessment question”:

*‘Are the agency’s evaluation function and its products: independent; credible; and useful for learning and accountability purposes, as assessed by a panel of professional evaluation peers against international standards and the evidence base.’*

34. As reflected in this question, the approach and methodology hinges on using recognized normative frameworks and **clustering** the treatment of the many issues **under three crucial criteria for evaluation**, specified as:
- A. **Independence** of evaluations and evaluation systems. The evaluation process should be impartial and independent in its function from the process concerned with the policy making, the delivery, and the management of assistance. A requisite measure of independence of the evaluation function is a recognised pre-condition for credibility, validity and usefulness. At the same time, each review should bear in mind in that the appropriate guarantees of the necessary independence in a particular organization will differ according to the nature of its work, its governance and decision-making arrangements, and other factors. Moreover, most organizations aim to encourage the active application and use of evaluations at all levels of management, meaning that systemic measures for ensuring the necessary objectivity and impartiality of this work should receive due attention.
  - B. **Credibility** of evaluations. The credibility of evaluation depends on the expertise and independence of the evaluators and the degree of transparency of the evaluation process. Credibility requires that evaluations should report successes as well as failures. Recipient countries should, as a rule, fully participate in evaluation in order to promote credibility and commitment. Whether and how the organization’s approach to evaluation fosters

partnership and helps build ownership and capacity in developing countries merits attention as a major theme.

- C. **Utility** of evaluations. To have an impact on decision-making, evaluation findings must be perceived as relevant and useful and be presented in a clear and concise way. They should fully reflect the different interests and needs of the many parties involved in development co-operation. Importantly, each review should bear in mind that ensuring the utility of evaluations is only partly under the control of evaluators. It is also critically a function of the interest of managers, and member countries through their participation on governing bodies, in commissioning, receiving and using evaluations.

### **Normative Framework**

35. In order to test the organization's evaluation function against these three criteria in the necessary depth, accepted normative frameworks are required as instruments for applying the assessment. Depending on their applicability to the organization being reviewed, one of several sets of accepted and consistent standards should be applied to provide such a normative framework:

- For UN organizations, the Norms and Standards produced in 2005 by the UN Evaluation Group (UNEG);
- For the participating Multilateral Development Banks, the good practice standards and template for independence of their Evaluation Cooperation Group; and
- For bilateral organizations, standards found in the OECD/DAC Principles for Evaluation.

The assessment against these standards can be supplemented by other questions judged relevant by the peer Panel (which should then be clearly specified in its report).

36. In order to proceed with applying the relevant norms and standards and ensuring proper coverage of the three main aspects, the norms and standards need to be broadly re-grouped and checked under those three headings. In the pilot studies, this process was carried out for the UNEG Norms and Standards and, in spite of some overlaps and repetitions<sup>4</sup> - inevitable with any such type of categorization - they did ultimately provide an adequate basis for organizing each review, its report, and the communication of its results. Similar work was done for the DAC Principles. Even taking account of the difficulties, the clustering of these norms under sharp and understandable headings such as the "independence, credibility and use" of evaluations clearly also helps in providing some thematic shape and comparability at an appropriate level. To avoid future professional peer reviews of UN organizations having to repeat this sorting and

---

<sup>4</sup> The UNICEF pilot peer review report was explicitly critical of the difficulties caused by the lack of a neat "fit" between the UNEG Norms and Standards and the categories of independence, credibility and usefulness, and recommended the search for a neater assessment approach, as well as some streamlining of the Norms and Standards. However, the Joint Task Force has not yet found any alternative assessment framework that would capture both the accepted standing of the Norms and Standards, and the useful clustering under these three important headings. It has also concluded from the two pilots that the difficulties are not insuperable and will ease somewhat in subsequent reviews. The goal of making suggestions to streamline the various sets of norms and standards in the field is a shared one, although it will require some more testing of their use and, of course, not just for peer reviews.

categorization work needlessly, such review teams are referred to the documentation of the UNDP and UNICEF peer reviews to be able to extract readily the relevant materials.<sup>5</sup>

37. The experience of the pilot reviews suggests that it may normally be appropriate to apply a combination of indirect (evaluation system) and direct (sample evaluation product) assessments at different levels. Although there are not yet single, internationally accepted standards for assessing the quality of evaluation products, a number of useful models are already being tested and applied at different levels. For example, a good number of the UNEG standards bear directly on the quality of evaluation reports, and in March 2006, the DAC Evaluation Network released a set of DAC Evaluation Quality Standards (on processes and products) for a 3-year test phase application. A number of organizations have also been developing and applying their own systems in these areas. Peer review teams and their host organizations should thus consider and decide at an early stage on the quality standards that should be applied, in an experimental spirit, for the review in question, and then document and report on the experience with their use.

38. In the same experimental vein, it should also be recognised that these peer reviews may be one of the most important continuing general tests of the usefulness of the prevailing norms and standards for evaluation, and they should thus be prepared to suggest ways to strengthen the norms on the basis of these experiences.

39. Given the central importance of the Normative Framework to the whole professional peer review process and its results, experience confirms the importance of the specific framework being agreed between the peer review team and organization at an early stage. As the *Lessons* from the UNDP pilot review stressed, "...future reviews should begin with a discussion and clear agreement by all parties on a brief outline of the framework and methodology. This should serve to clarify expectations and provide a sound basis for agreement on a more detailed review plan and work-programme."

### **3. Roles and responsibilities in the reviews**

#### *The Peer Panel*

40. The expertise and commitment of the peer panel itself, together with the cooperation of the organization being reviewed, is essential to the successful conduct of the peer review. Given the likely demands of other responsibilities on most panel members, the demands of this work need to be well-organized and prepared, and highly experienced and skilled consultants are also likely to be required.

41. In selecting participants for the panel, which should be restricted to five or six participants, experience confirms that the major criteria for selection of the group should include a combination of the following attributes:

- i. Independence from the particular organization being assessed;
- ii. Professional evaluation expertise;

---

<sup>5</sup> In summary, for UN organizations, indicators of independence are broadly covered by UNEG Norms N6.1 – N6.5 and amplified in the relevant Standards. Indicators of credibility are mainly treated in UNEG Norms N5.1 – N5.3, N8.1, N9.1 – N9.3 and N11.1 – N 11.5 and amplified in the relevant Standards. Indicators of utility are mainly treated in UNEG Norms N2.6, N1.3, N 8.2, N10.1, N 10.2 and N.12.1 – N12.3 and amplified in the relevant Standards

- iii. Understanding of the context and use of evaluation in development cooperation and multilateral organizations;
- iv. An acceptable gender mix;
- v. Participation from multilateral and bilateral agency/ies;
- vi. Participation from country/ies receiving assistance, including those with evaluation responsibilities;
- vii. Participation from independent evaluation experts; other research fields, oversight disciplines or knowledge sharing expertise; and
- viii. Capacity to deal with senior management and governing body levels.

42. It is recommended that either one bilateral evaluation office represented on the peer panel or several members of the peer panel take the **lead responsibility** for:

- Managing the process and either directly providing logistical and secretariat services, or contracting in relevant support;
- Being willing and ready, through representation on the multilateral organization's board, to engage with the agency's senior management and governing board to actively encourage their interest, and the consideration of recommendations and possible action following the assessment process; and
- Ensuring that the key conclusions and possible lessons from the review are communicated to the wider UN Evaluation Group and DAC Evaluation Network membership and that the reports and supporting evidence are easily accessible to members and others.

43. Although the peer panel takes overall responsibility for the process, together with the organization reviewed, the most intensive direct roles of panel members would normally be concentrated near the beginning and end of the overall process:

- First, in clearly defining the framework and *modus operandi* to be proposed for the particular review, with the organization being reviewed<sup>6</sup> (phases 1-3 below); and
- Later in validating the evidence and findings, bringing together the judgements, conclusions and recommendations of the review, presenting and following up the results. (Phases 5 & 6 below).

The major part of the intervening data-gathering and analysis work (phase 4 below) would normally be expected to be delegated to carefully-selected consultants,<sup>7</sup> responsible to the panel and working within the agreed framework.

---

<sup>6</sup> To ensure mutual clarity and smooth working relationships, the panel and its consultants should normally take responsibility for briefly setting out the agreed plan for the specific peer review on paper for reference as required. It should reflect the application of this framework and specific adaptations.

*The evaluation unit of the organization under review*

44. The evaluation unit is likely to be the internal initiator, principal organizer, and substantive collaborator on the review within the organization concerned, although to the extent that a governing body or senior management are willing and able to take more active roles in the first two areas this is likely to have major benefits to the credibility and usefulness of the review. The evaluation unit (or other principal partner to the review from the organization) will be undertaking the responsibility for considerable work in the following areas:

- Informing and obtaining the necessary endorsements by senior management, and if possible the governing body, to conduct the review and assure the necessary cooperation;
- Up-dating or preparing an initial self-assessment of the evaluation function in the organization;
- Considering and agreeing to a specific design and plan for the review;
- Facilitating all necessary access to its own data and internal interviewees, and assisting in identifying and securing similar access throughout the rest of the organization;
- Helping ensure the presentation of the results to senior management and the governing body, and its dissemination throughout the organization; and
- Contributing its views on the experience to the Joint Task Force, and providing follow-up reports (after one and two years) on actions taken and other impacts of the review.

**4. Phase-by-phase organization of a professional peer review process**

(illustrated with typical steps and sequences).

**Phase 1: Initiating and organizing the review.**

- i. Invitation by the organization (preferably the governing body or senior management) to conduct the professional peer review. It should be noted that such peer reviews are likely to be best planned and scheduled to mesh with important decisions, reviews and/or relevant events;

---

<sup>7</sup> The 2005 approach and methodology for the two pilot studies had also envisaged the options of panel members or their colleagues carrying out this work or, if agreed, the host organization itself. In the event, two-member teams of consultants from different countries were retained to carry out this work. In addition to a high level of skills in evaluation, synthesis and communication, the selection of consultants should also reflect key requisites for the panel, such as independence from the particular organization being assessed, a very good understanding of the context and use of evaluation in development cooperation and multilateral organizations, and the ability to work well in this capacity with the host organization at various levels as well as the panel. Including consultants from country/ies receiving development cooperation is an important objective. The two pilot studies had one two-male team and one two-female team – whether or not a gender mix is always possible, good knowledge of gender issues should be a requirement on the consultant team. It is important that the consultant team be in place early in the process to acquire an excellent grasp of the panel's approach and the arrangements with the host organization.

- ii. Volunteering (ideally in some rotational order) by a representative of a lead peer organization, willing to offer direction, coordination and financial support, backed by a consortium of other peer organizations also willing to participate. (In addition to the substantial in-kind contributions required of them, some organizations being reviewed might in some cases want to take some share in the direct costs.);
- iii. Endorsement by senior management of the host organization and if possible by its governing body. (At the very least, the latter should be informed and engaged at the outset and at key milestones in the process. Demonstrating benefits from other experiences may be helpful in stimulating interest and demand by both senior management and governing bodies);
- iv. Constitution of the peer panel by the lead organization and consortium members, in consultation with the host organization concerned and the Joint DAC/UNEG Task Force;
- v. Selection and appointment of consultant/s to support the peer panel;
- vi. Preparation and circulation of some form of self-assessment of its evaluation function by the organization concerned (e.g., for UN organizations, up-dating of the responses to the 2005 UNEG self-assessment checklist, or an appropriate equivalent for other organizations);
- vii. Thorough familiarization, normally by the review consultants, with the evaluation function in the context of the organization's mission, governance, policies, structures, budget, staffing and operations. It is important to first try to understand the organization, and how the evaluation function can best serve it, prior to looking in depth at the machinery.

**Phase 2: Agreeing on the review methodology, work-program and mutual responsibilities with the host organization.**

- i. The panel would normally receive intensive orientation briefing by its consultants on the organization and its work, and the results of its self-assessment of the evaluation function. This could be followed by direct meetings with the evaluation unit and possibly others to cross-check and verify this starting base;
- ii. Following the previous step, the peer panel should pursue a discussion and clear agreement with the evaluation unit on the core assessment question, the basic normative framework and methodology, the work-program and mutual responsibilities for the conduct of the review, including an outline of the succeeding phases, specifying anticipated data requirements, scheduling and arrangements for access to sources, interviewees, possible field missions, etc.
- iii. The panel provides guidance to consultants on its priority interests for the collection and analysis of data, and preparation of findings. Depending on the character of the organization's evaluation functions, data-gathering and findings might be guided by emphasis on centralized or de-centralized evaluation processes, appropriate and feasible

samples of individual evaluations and their products, etc.<sup>8</sup> An “issues paper” could be the appropriate form to gather the evidence and analysis in this stage of the process.

**Phase 3: Collecting and analysing the data, shaping findings against the normative framework adopted**

- i. Normally carried out in the main by the panel’s consultants, in consultation with the evaluation unit and others, this phase would be likely to include: extensive review of the relevant files and documentary information base to be gathered and analysed in line with the core assessment question and the normative framework; key informant interviews, focus groups and field missions around systems and processes and/or particular evaluations, which are used as reference cases;
- ii. Preparation of the base of evidence and findings for the panel’s review.

**Phase 4: Validating the evidence and findings, shaping conclusions and recommendations.**

- i. Peer interviews: With the benefit of the information assembled, and its individual and collective examination by the panel, including selected primary sources, and any observations from the evaluation unit, the Panel will conduct wider interviews with a number of Senior Managers and Executive Board members concerned. The reference cases will serve as one important focus;
- ii. Integrating the results and insights from these interviews, the Panel will complete the triangulation, refinement and confirmation of its base of evidence and findings. On that basis, it will then move into the “judgement phase” following the following steps: panel agrees on its main frameworks for judgements in relation to the Norms and main aspects of the assessment; panel considers and debates the evidence and findings, and arrives at its draft conclusions and recommendations, agrees on draft report;

**Phase 5: Presenting and discussing the review report, dissemination**

- i. Draft report serves as the basis for the Peer Assessment meeting with the host evaluation unit to consider the results in depth;
- ii. Panel carries out final review and any necessary revision to the draft report;
- iii. Panel transmits its final report to the host evaluation unit, providing a final opportunity for review and reflecting dissenting views or other key responses;
- iv. Panel presents the Peer Review Report to the organization, ideally at the Governing Body and accompanied by a Senior Management Response.

---

<sup>8</sup> The two pilot reviews illustrated the need to adapt the approach. The review panel for UNDP focused on its centralized evaluation function and office. To make the informational foundations as concrete, focused and systemic as possible within the scope of the assessment, it selected a small group of recent evaluation processes and products of different types carried out by the Evaluation Office as reference cases, taking into account a combination of several criteria. In the case of UNICEF, with a highly de-centralized evaluation system, the methodology placed considerable emphasis on an intensive examination of evaluation within a particular country program, Ghana in active engagement with Ghanaian and other partners in that country.

## **Phase 6: Following up**

45. Panel and organization reviewed report, ideally together, to the Joint UNEG DAC Task Force on lessons learned from the exercise, and its results over time.

### **5. Time and resources required**

46. There will always be too many variables to allow any firm across-the-board estimates in advance for the time and resources needed for professional peer reviews of evaluation in different organizations (e.g. size and complexity of the organization of the organization and its programs; locations of panel members; calendars of work for the host organization and panel members, amounts of field work required, etc.)

47. At the same time, it may be helpful - on the basis of experience to date and discounting the up-front “development costs” of the first two “pilots”- to provide at least *a general sense of the minimum time and resources* likely to be needed to carry out a peer review along the lines outlined in this document. The general estimate of elapsed time required would be expected to be at least six months from inception to reporting, but should not be more than nine months. Over such a period, panel members would each be required to devote at least 10-15 days of work, and the lead organization/s at least an additional 10-15 days of supporting and logistical work. The panel’s strong supporting advisors or consultants would probably require a total of at least 60 days work. In the host organization, meanwhile, the evaluation unit will probably be required to devote at least 30-35 days of professional time, and significant administrative facilitation, while their colleagues elsewhere in the organization will be called upon for the time for interviews, assisting with field missions, etc.

### **6. Planned monitoring and assessment of experience**

The Joint Task Force DAC/UNEG will continue to work on the following:

1. conducting a quick joint review of experiences following each peer review – as standard procedure;
2. stocktaking once every two years on the basis of such reviews;
3. updating of the approach and methodology (described in the current framework) on the basis of this periodic stocktaking.

### ***Key References:***

1. Peer reviews of evaluation in international organizations, A joint initiative of the DAC Evaluation Network and the UN Evaluation Group, Draft August 9, 2006
2. Peer Review of Evaluation Function at UNICEF, Peer Panel Observations and Recommendations, June 22, 2006
3. Peer Review of UNDP Evaluation Office, Key Issues and Lessons Identified, 17 March, 2006
4. DAC Evaluation Network, A New Approach to Assessing Multilateral Organisations’ Evaluation Performance, Final Draft. June, 2005

5. Peer Review: United Nations Development Programme Evaluation Office. December, 2005, DANIDA et al.
6. Peer Review of the Evaluation Function at United Nations Children's Fund (UNICEF). May, 2006
7. UNEG Norms and Standards
8. DAC Principles for Evaluation, and up-date
9. TORs for Drafting a Framework for Future Reviews