Lessons-Learned Study
of
Peer Reviews of UNEG Evaluation Functions

Final report

Ian C Davies & Julia Brümmer

May 2013

Preface

We take great pleasure in prefacing this report of the Lessons-Learned Study of Peer Reviews of UNEG Evaluation Functions.

Under the auspices of the Joint UNEG/EVALNET Task Force on External Peer Reviews, the first review was undertaken in 2005. Since then, ten peer reviews of evaluation functions of UN agencies have been successfully carried out.

Each one has brought its share of valuable lessons for improving the quality of the agency's evaluation function and its contribution to agency performance. In this way the reviews are intended to provide assurance to donors funding the agencies that they can rely on the agency's own evaluations.

Since 2005 the peer review process itself has been refined and enhanced by making it more adaptable to different organisations and contexts. In particular peer reviews have sought to reap fully the benefits of sharing expertise from different agencies with different roles.

The report summarises a stock taking exercise which marks a coming of age for the peer review process. It is now recognised as a critical component of building assurance about the quality of evaluation and the performance of agencies.

Furthermore, because of their multifaceted value, peer reviews are increasingly sought after by communities of evaluation to strengthen theory and practise and to meet high standards of professionalism.

The sustained cooperation between UNEG and EVALNET is a model of professional collaboration that we are proud of and without which the accomplishments of the peer review process would not have been fully realised.

We trust you will find this report a useful reference as we strive collectively for meaningful high quality evaluation in our respective fields of practice and organisational settings.

Deborah Rugg, Chair   Martyn Pennington, Chair
United Nations Evaluation Group (UNEG)   DAC Network on Development
                                              Evaluation (EVALNET)

Table of contents

1. Introduction

This report presents the findings, analysis, conclusions and recommendations of the study of peer reviews of UN evaluation functions, which was conducted from mid-January to end of March 2013.

2. Background

As part of the overall trend towards greater accountability on development effectiveness, the OECD/DAC Network on Development Evaluation planned to establish a peer review process of the multilateral evaluation functions in an attempt to reduce costly institutional evaluations funded by donors. The DAC Network on Development Evaluation (EVALNET) and the UN Evaluation Group (UNEG) established subsequently a Joint Task Force to initiate and support professional peer reviews of the evaluation function of UN organizations. For that purpose, a framework for professional peer reviews was developed in 2006, based on internationally recognized standards and two pilot reviews of UNDP and UNICEF. The reviews are conducted by an independent Peer Panel consisting of professional evaluators with the support of expert advisors. Lessons from earlier peer reviews are incorporated in the approach, and the framework was revised in 2010 to become a fully owned UNEG framework for peer reviews[1]. This framework can be adapted according to the context of the specific organization whose evaluation function will be reviewed.

As stated in the framework, the peer review has several purposes, including: (i) the provision of an independent and professional assessment of the extent to which the UN Norms and Standards for evaluation have been adopted in the organization concerned; (ii) building greater knowledge, confidence and use of evaluation systems by management, governing bodies and others; (iii) within the organization, providing a way of "evaluating the evaluators" through a professional peer assessment against the UN Norms and Standards; and (iv) sharing good practice, experience and mutual learning, building greater internal capacity and external confidence in UN evaluation systems.

The peer review assesses the extent to which the evaluation function contributes to accountability and learning in the organization for: the governing body and senior management of the organization, its member states, and communities and donors providing support to the organization. Furthermore, the peer review also provides a perspective on the extent to which evaluations cover the achievements of the organization.

---

[1] The UNEG Framework for Professional Peer Reviews of the Evaluation Function of UN organizations was approved as a UNEG reference document at the UNEG Annual General Meeting 2011.

The peer review approach was piloted in assessing the evaluation function of UNDP in 2005 and of UNICEF in 2006. Since then, peer reviews were undertaken of the evaluation function of WFP (reported in 2007), OIOS (reported in 2009), GEF (reported in 2009) and UNIDO (reported in 2010). The reviews of the evaluation functions of UNEP, UN-Habitat and FAO were reported in 2012. In coming years peer reviews will be undertaken of the evaluation functions of other UN agencies.

Over the past 8 years, significant experience has been built and lessons have been drawn from each peer review. There is a continuing interest in the international community for greater accountability on aid effectiveness as reflected in the various modalities developed in the past few years for evaluating multilateral organizations. At the same time, there is a continuing pressure for strengthening evaluation functions and for increased harmonization of the evaluation functions, in particular, within the UNEG. It is timely to consolidate lessons from this wealth of experience and assess the extent to which the peer reviews have fulfilled their various purposes.[2]

3. Objective of the study

The objective of the lessons-learned exercise is to assess whether (or not) (a) the reviews have made a difference in the use made of evaluations undertaken by UN evaluation departments by different stakeholders, in particular by donors; and (b) whether these reviews have had a positive impact on the evaluation departments of the UN agencies which were subjected to such reviews. In addition, the study aims to provide insights on the usefulness and feasibility of the Peer Review approach as outlined in the UNEG Framework for Professional Peer Reviews. [3]

4. Methodology

Nine peer reviews were examined for this study which pursued three lines of inquiry. The study team conducted jointly:

- A structured review of all significant peer review documentation including peer review reports, management responses, lessons learned documents, notes, norms and standards and frameworks. As well the study team reviewed documentation related to the professionalisation of evaluation, to peer review in comparable professions as well as to quality control and assurance systems and practises.

- Semi-structured interviews of about an hour each by telephone and by Skype with twenty one peer review panel participants covering the nine first peer reviews that have been conducted.

---

[2] From the terms of reference
[3] From the terms of reference

•A web based survey of bilateral evaluation and multilateral departments.

5. Findings

5.1 Quality of processes

The view was expressed by interview respondents that, in the initial peer reviews, the first version of the framework presented the process more as an evaluation than as a peer review. In one case the evaluation function that was being reviewed considered that it spent a significant amount of time explaining what it did and that the peer exchange element was missing.

A number of interview respondents who had participated in peer reviews considered the process too lengthy, i.e. spread out over too long a period of time.

The majority of interview respondents from other UN agencies and bilateral organisations and who were members of peer review panels consistently highlighted the value of learning about other evaluation functions, their practises and challenges, and of engaging in constructive exchanges with peers. *"The great strength of the process is that you get to know each other and you get to know different organisations. The strength lies in institution building and for building bridges."* This return on investment appeared to offset to a significant extent what these members assessed generally as being a demanding process in terms of time commitment.

Generally panel members that were interviewed found the involvement of an advisor to the review panel, i.e. an external consultant, useful and necessary.

Team composition is considered a key factor for a well performing review panel. A number of interview respondents stressed the importance of selecting members with appropriate evaluation expertise and experience. An important lesson was that the panel team should be guided by a person with very strong evaluation knowledge.

According to a respondent, one of the challenges is that panel members tend to come to the peer reviews with the assumption that the way their function works is how it works everywhere, however UN evaluation functions are very different. In these situations external advisors who are consultants may bring considerable value added in as much as they are often familiar with a number of different organisations and evaluation functions.

Respondents noted that the timing of a peer review is also a key factor in its success. In a number of cases peer reviews were requested by the newly arrived head of the evaluation function as a means, among other things, to get the time of day on its state. *"The right time to do a peer review is when people are ready to listen."*

5.2 Adherence to framework

Adherence of peer reviews to the UNEG framework was assessed through a review of the nine peer review reports. For this purpose a checklist was developed including the most important criteria in terms of peer review process and content laid out in the UNEG framework. A summary checklist including the combined results for all nine reports is included in annex.

Peer review reports overall show evidence of good adherence to the peer review framework – independently from whether they were produced before or after the development of the first and second version of the framework. This could be seen as a confirmation of some of the statements from the lessons learned and interviews that the framework is helpful as a starting point to organise the peer review process rather than as a tool that can easily and strictly be adhered to.

Of the nine peer reviews, six seemed to follow the "full-fledged" process of peer reviews while three seemed to adhere to the reduced version of the process.[4] Most of the reports do not explicitly spell out which version was adhered to.

In terms of process, only the criterion of "seriously engaging developing partners" in the process was not adhered to by most peer reviews. As for the selection criteria for the panel members, almost half (four out of nine) peer reviews followed these criteria only partly or not at all, and only part of the reports where the panel composition did fit the selection criteria set out in the framework made actual reference to them. It is therefore unclear whether they deliberately followed the selection criteria.

In terms of peer review content, most of the items in the framework were followed across the peer review reports. The two items that generally received less attention were the "coverage of organisational achievements" (often touched upon in the reports, but not systematically assessed and remaining without any clear conclusion) and the "collaboration with local partners and stakeholders as well as harmonization and coordination with other external partners", which is only assessed explicitly in two of the peer review reports.

---

[4] In line with the "reduced" approach laid out in the 2011 UNEG Framework for Professional Peer reviews, the main criteria used for defining a peer review as "light" was the integration of phases 2 (fact-finding) and 3 (peer exchange) into one single visit by the panel to the peer reviewed organisation. This was the case for the UNIDO, UNEP, and UN-Habitat peer reviews. Even though the UNEG Framework assumes that in a reduced peer review, no consultants will be hired all of these three peer reviews were carried out with the contribution of an advisor.

5.3 Usefulness of frameworks

The framework was found useful generally as a starting point for structuring the peer review planning discussion and for providing general guidance. Interview respondents emphasised however that the framework was used and adapted based on the particular circumstances of the evaluation function under review and of the constraints and limitations of time and resources.

Some respondents stated that they found the framework somewhat rigid and prescriptive, with a very strong focus on norms and standards that did not always fit well with the evaluation function under review and its context. These remarks tended to apply however to the first version of the framework.

The framework was considered useful generally in setting out a process and structure for the peer review although some respondents found that it gave little against which to judge the performance of the evaluation function.

"The UNEG framework does not necessarily get the important information out. There is actually no need for a rigorous framework. The framework was never really "used", but it was useful to organize the work in the beginning."

A couple of respondents expressed the view that, although the framework was found useful, some of the norms and standards that it is anchored in lack clarity and should be updated to reflect developments in evaluation theory and practise.

5.5.1   Reduced peer reviews

According to interview respondents the current updated framework gives the option of reduced peer reviews and is considered generally to be more flexible, more adaptable and hence its application can be less costly than its previous version. For some respondents the updates to the framework also reflect a desire to allow for more peer exchange as part of the peer review process.

Respondents defined reduced peer reviews as those that did not include field visits and for which there could be reduced scope such as the number of partners selected for interviews. However respondents stressed that the key elements of light or reduced peer reviews were not substantially different from the full peer review.

The respondents that were interviewed considered that all of the key elements of the framework were addressed as part of the peer review process they were involved in although to different degrees of depth. As such, they were of the view that the framework had been well respected in its application and use.

"Even "light" peer reviews are cumbersome; having a lean process is a bit tricky as 1. UNEG norms and standards are complicated, 2. Need to substantiate & triangulate, 3. Need to get to know the organisation."

5.4 Quality of reports

The study's review and assessment of the quality of the peer review reports[5] matched a strong consensus among interview respondents that rated as very good overall the quality of the reports. *"The quality of most peer review reports is pretty good. Most elements of the framework are covered, they are quite readable although not all aspects are always treated in depth."*

Although overall the quality of reports is considered good, it was pointed out that these are not the result of an evaluation so that sometimes the ways in conclusions are reached give the reader a sense that these are based more on perceptions and individual judgments than on factual evidence.

It was also noted that the greater part of the peer review reports is generally descriptive. As well, respondents noted that data was overwhelmingly qualitative and that this sometimes affected the perception of credibility stakeholders had in the robustness of peer reviews.

"Overall the reports are quite good in that they are readable and understandable however in terms of content it is difficult for an outside reader to know what's not there and therefore to judge their quality."

5.5 Usefulness against purposes

"The time has come to reconsider fundamentally the role and purpose of peer reviews"

5.5.1   Use of UN evaluations by donors

A majority of interview respondents were doubtful that the peer reviews provided assurance to bilateral donors such that they placed increased reliance on the work of the peer reviewed evaluation functions and diminished their evaluation of the corresponding agency's programmes accordingly.

Different possible explanations were put forward:

• Within bilateral organisations the individuals who deal with multilateral organisations such as UN agencies are not necessarily associated with the peer review process and are not the ones who participate in the EVALNET-UNEG meetings. As one respondent put it: *"Those who commission evaluations of multilateral agencies are not at the table".*

• Although peer reviews might be regarded in a positive light by bilateral organisations there is still a view that the mechanism is "evaluators assessing evaluators" which may diminish the perceived credibility of the exercise and the degree of reliance donors may place on the evaluation function and its work.

---

[5] Please see in annex : Quality of peer review reports

- Bilateral agencies with responsibility for development assistance have their own obligations to provide evaluative information that is relevant to their national political and policy processes as well as appropriate accountability information to their Parliaments and stakeholders. As such, they, together with their evaluation functions, are mindful of the need to ensure the credibility of their representations as well as the relevance of evaluation choices to domestic political contexts.

There was an expressed sense among an appreciable number of interview respondents that donors *"were stepping back from the peer review process"* although it was not fully clear why. Some suggested that donors had more of an expectation that peer reviews would provide assessments of the performance of UN agencies along the lines of The Multilateral Organisation Performance Assessment Network (MOPAN).

The view was expressed that so few peer reviews are carried out of evaluation functions that it is not an adequate tool to provide reliance to donors.

Most respondents were of the opinion that there was a leap of logic in associating peer reviews with improved organisational effectiveness.

A number of respondents considered that peer reviews were not very cost-effective and that there might be better ways of assessing the effectiveness of multilateral evaluation functions.

5.5.2   Impact on evaluation functions

Respondents overall were of the view that peer reviews had a positive impact on the evaluation departments of the reviewed UN agencies. Respondents were unanimous in stating that peer reviews had increased the profile, credibility and value of the evaluation function in the organisation. *"Peer reviews have achieved the goal of increasing the independence of evaluation functions and the adoption of new evaluation policies."*

Recommendations that more strategic level evaluations be conducted were generally positively received by governing bodies of peer reviewed agencies and considered by respondents to be a positive impact on both the evaluation function and the organisation.

The view was expressed by respondents that a key target of peer reviews is the internal management of the organisation and that they should not be an instrument for external partners and donors.

"There is little value in assessing an evaluation function as good or excellent if the organisation is not performing or being managed well"

5.6 Implementation of recommendations

"We presented recommendations to the Board but then our job was done. So we don't know what happened."

Recommendations can broadly be grouped under the three criteria of independence, credibility and usefulness (some of the recommendations obviously relate to more than one criterion). There is a wide range of recommendations and the level of detail varies from one report to the other.

Some of the recurrent recommendations are on the following issues:

Independence

- Reporting lines of the evaluation function and content of reporting
- Formalisation of responsibilities and procedures in an evaluation policy
- Recruitment of external experts avoiding conflicts of interest
- Ensuring independence in programming, budgeting and planning

Credibility

- Quality assurance of evaluations (some recommendations point at details regarding content or methodology of evaluations)
- Adequate budget to enable quality evaluations
- Capacity strengthening of human resources in the evaluation function (training, upgrading of staff positions, adequate number of staff)
- Evaluation functions to take final responsibility for contents of evaluation reports

Usefulness

- Use of evaluations for strategic decision making
- Responsibilities and process for management response to evaluations
- Relationship of evaluation with the RBM system of the organisation (most recommendations are on addressing existing fragmentation of the RBM system)
- Systematic harvesting of lessons learned, dissemination of findings
- Involvement of stakeholders throughout the evaluation process, starting from planning

Only one report (GEF peer review) contains lessons learned and some of these actually do not really fit the definition of lessons learned.

5.7 Overall strengths and weaknesses

Strengths

Peer learning and peer exchange are considered key strengths of the peer review process. In particular, the peer review process provides the opportunity for evaluators from different agencies to learn about each other's realities and practises.

As well, raising the profile and highlighting the value of the evaluation function, especially at the policy and strategic levels, within its organisation is seen as an important value added of the peer review mechanism.

The peer review process is seen to increase the trust that bilateral donors have in the evaluation functions that have been peer reviewed and the accountability of the UN agencies of which they form a part.

The peer review reports are considered by all involved to be of high quality and useful to the peer reviewed evaluation functions and their agencies.

Weaknesses

The view was expressed that one weakness of the peer review mechanism is that it is not mandatory. One of the results being that the stronger evaluation functions tend to be the ones that volunteer for, or request, a peer review.

A few interview respondents considered that a weakness of the process overall was that once the peer review was over, panel members did not systematically receive feedback at a later date on what may have changed as a result of the peer review and of its recommendations. *"The big question is do the evaluation departments of the multilaterals use the peer review report and change for the better?"*

For a number of respondents there could be a potential for bias in how the peer reviews are resourced, if the donor financing the peer review has a predominant say in the composition of the peer review team and the selection of external advisors.[6]

As well, a number of respondents found that the review process did not focus enough on learning but rather on exercising judgment about the quality of the evaluation function and gathering required evidence.

---

[6] Whereas this may have been a risk in earlier peer reviews, the present procedure is that the head of the panel is usually approached by the EVALNET-UNEG task team and the panel and advisor composition is decided in close consultation with the head of the evaluation function of the peer reviewed organisation and in order to avoid possible conflicts of interest.

Respondents stated fairly consistently that, in their view, the peer review process could be unnecessarily lengthy and costly, and that budget considerations could drive process choices rather than the purported purpose of the review.

5.8 Other

Of note is the peer exchange workshop that took place with staff from FAO toward the end of that peer review process. According to respondents who participated the meeting, which took place on a Saturday, was well attended by staff of the evaluation unit and the discussion was rich and highly appreciated. The experience appears to stand out as a highlight and as well as a "lesson learned" about the importance and value of staff involvement and of peer exchange.

6. Summary of survey results

The study team conducted a web-based survey of bilateral donors. The survey was pre-tested in mid-February 2013 and sent out to 28 members of EVALNET in the last week of February. These EVALNET members were asked to complete the web-based survey and to send the survey link to their colleagues in their multilateral departments, in particular those individuals dealing with the UN agencies that had been peer reviewed. To the knowledge of the study team, the survey in total reached 59 people and 23 responded, for a response rate of 39%.

Most of the respondents (18) are working for a national ministry (mostly foreign ministry or development department); at least nine different countries were represented.[7] About half of the respondents (12 out of 22 answering that question) are working for the evaluation function of their ministry or organisation, eight are working for the department for multilateral or UN agency relations, 1 for senior management and 1 for another department.

Correspondingly, most respondents rated their knowledge of evaluation as "very good" (27%) or "excellent" (41%). A majority of respondents (65%) said they were "somewhat" familiar with the peer review process, 26% said they were very familiar, 9% (2 respondents) said they were not at all familiar with the process.

6.1 Impact of peer reviews

Responses indicate a positive impact of the peer review process in terms of greater trust in the multilateral evaluation functions and greater reliance on evaluation reports from the peer reviewed agencies (see figures 1 and 2). One respondent stated that peer reviews *"[...] are playing an important role in influencing the quality and credibility of the evaluation departments as well as a tool for mutual learning between the reviewed evaluation department and those in the review itself."*

---

[7] 15 respondents indicated their country of origin. They are from the Netherlands (4 respondents), Switzerland (3 respondents), Denmark (2 respondents), Austria, Luxembourg, Norway, Spain, United Kingdom, and the United States (1 respondent each)
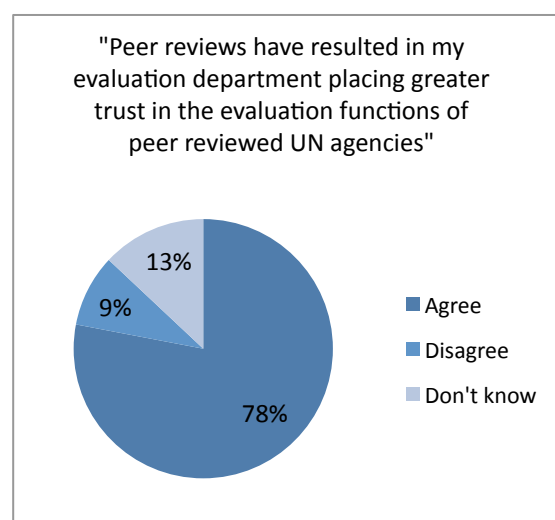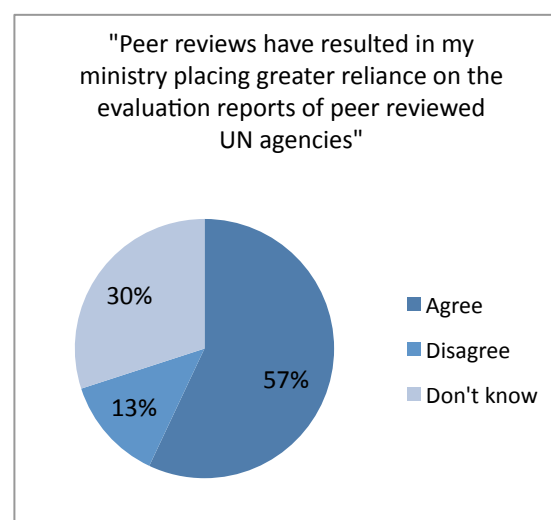
Figure 1

"Peer reviews have resulted in my evaluation department placing greater trust in the evaluation functions of peer reviewed UN agencies"

13%
9%
78%

- Agree
- Disagree
- Don't know

Figure 2

"Peer reviews have resulted in my ministry placing greater reliance on the evaluation reports of peer reviewed UN agencies"

30%
57%
13%

- Agree
- Disagree
- Don't know

The picture is less clear when it comes to closer working relationships between ministries and peer reviewed agencies: 35% believe this has been the case, 22% believe it has not, almost half of the respondents do not know.

Most of the respondents agree that peer reviews have led to improvements in the accountability of peer reviewed UN agencies (figure 3) as well as in the quality of evaluation reports of peer reviewed UN agencies (figure 4). At the same time, the majority of respondents believe that the peer reviews have not led to a reduction in the number of evaluations their ministry does of programmes of peer reviewed UN agencies (figure 5).

Figure 3

"Peer reviews have led to improvements in the accountability of peer reviewed UN agencies"

5%
18%
14%
64%

- Strongly agree
- Agree
- Disagree
- Don't know

Figure 4

"Peer reviews have led to improvements in the quality of evaluation reports of peer reviewed UN agencies"

5%
45%
50%

- Strongly agree
- Agree
- Don't know

Figure 5

"Peer reviews have led to a reduction in the number of evaluations my ministry does of programmes of peer reviewed UN agencies"

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Don't know

5% · 24% · 24% · 4% · 43%
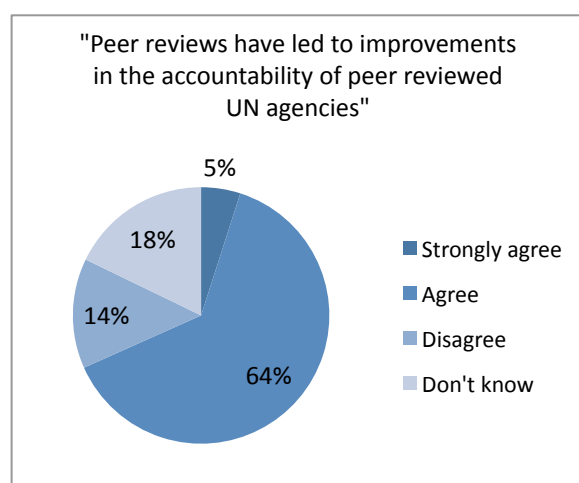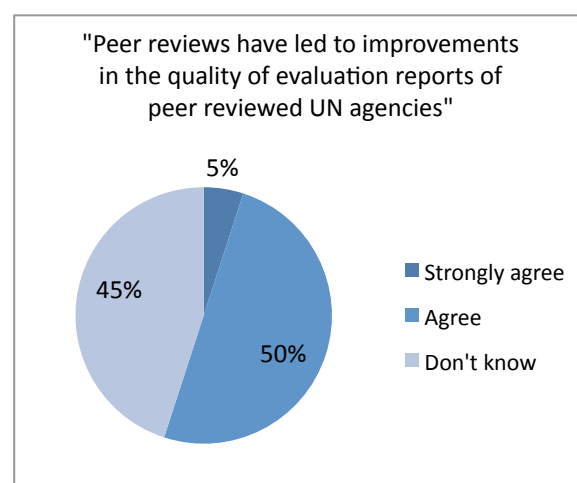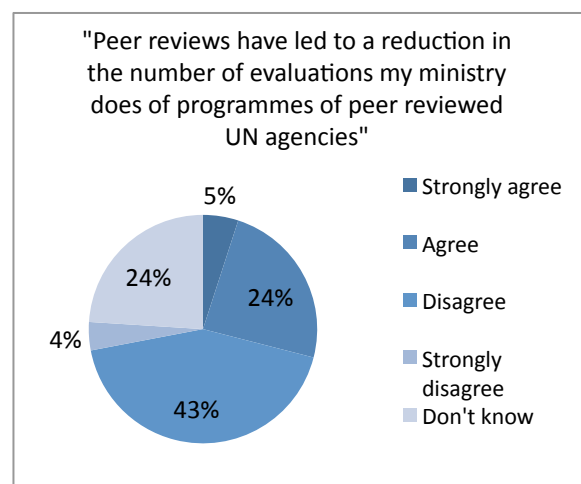
As part of their additional comments (free text response), three respondents expressed doubts about the extent to which peer review reports are sufficiently discussed and used by their ministries' officials in charge of the relations with UN agencies.

6.2 Quality and credibility of peer review reports

Most of the respondents judged the peer review reports to be of very good (9%) or good (59%) quality. No respondent found the reports poor or very poor.

Similarly, most respondents assessed the credibility of the reports as either very good (14%) or good (50%). Only one respondent found that the credibility was very poor (5%). One respondent explained that the credibility of peer reviews may suffer from the fact that evaluation and evaluators tend to set themselves apart from other management functions, which might lead to some suspicion by organisations.

6.3 Purpose of peer reviews and role of EVALNET-UNEG

Both the performance assessment and the learning and peer exchange aspects of the joint EVALNET-UNEG peer review process are seen as important by the majority of respondents. Most of the respondents either strongly agree (23%) or agree (55%) that the primary purpose of the peer reviews should be to assess the performance of UN evaluation functions; 14% disagree. The responses to the question whether the primary purpose of the peer reviews should be to foster peer exchange and learning for UN evaluation functions are very similar: 27% strongly agree, 55% agree, 14% disagree. No clear preference for either of the two purposes is visible from the survey results.

7. Summary of lessons learned

The study team reviewed lessons learned documents for seven out of the nine peer reviews (UNDP, WFP, OIOS, GEF, UNIDO, UN-Habitat, UNEP). These documents were mostly prepared by the peer review panels and in some cases also by the head of the peer reviewed evaluation function. They summarise relevant experiences from the peer review processes, usually in the form of informal feedback to the joint OECD-DAC/UNEG task force on peer reviews.

7.1 Concept of "peer review"

Particularly in the context of the WFP peer review, difficulties with the concept of "peer review" were highlighted. Finding the right balance between close collaboration with the peer reviewed organisation on one hand and maintaining the integrity and independence of the peer review panel on the other was seen as a challenge. This challenge implies a difficulty of keeping data collection and tentative analysis distinct from the actual review process, which is done together with the peer reviewed organisation. It was noted particularly that evaluators struggle with understanding and adopting a genuine "peer review approach" due to their evaluation background.

7.2 Peer review process

Some of the lessons learned documents highlight the importance of scheduling the peer review exercise keeping in mind important decisions, reviews and relevant events within and beyond the agency being assessed.

One document stresses the need to clearly establish at the outset how the forthcoming peer review will be handled within the reviewed organisation in order to fully understand the context for the peer review.

Issues of financing peer reviews were taken up in a few lessons learned documents. One opinion which is put forward is that UNEG (and EVALNET) should participate in the financing of peer reviews, first, to enhance the possibility of UNEG leading the process and taking full ownership and second, because financing a peer review panel member can be a heavy financial burden on organisations that only have small budgets for their evaluation function.

Clarity about the primary audience of the peer review report is seen as important and should be considered in the design of the peer review approach. One document suggests that a stakeholder analysis could be a useful component within the first part of the peer review process.

Field visits are found useful in the case of some peer reviews, e.g. WFP, while others conclude that their value-added should be carefully considered given the need for additional resources that they imply. One lessons learned document concludes that field visits should only be considered in the case of largely decentralized organizations and when the factual report or the interviews raise serious issues regarding impartiality or credibility of the evaluation office.

Reflecting on its experience, one peer review panel concludes that the peer reviews should put stronger emphasis on other informants outside the evaluation office, i.e. "users" of evaluation, to generate more evidence on the aspects of utility and credibility of evaluations produced by the reviewed function.

With regard to the actual "peer" process, one document suggests that this aspect of the process should be strengthened, to allow for sufficient time to discuss issues and their potential solution.

A few lessons learned documents highlight challenges in doing the data analysis and agreeing on the final report given that the panel has to do this at distance. It is proposed to allocate sufficient time for this part of the process, in order to sufficiently share and collectively digest the information that the various panel members and advisors have obtained during the data collection phase. One peer review panel concludes that a face-to-face meeting to discuss the draft report is preferable. This would however need to be budgeted for at the outset.

7.3 Peer review panel & advisors

Many of the lessons learned documents include comments and suggestions about the selection of panel members and advisors as well as their respective roles. One lesson learned is that the panel should be tailored to the organisation to be reviewed and members should be selected in a manner that reflects the organisation's specific profile, including those who are not necessarily staff of bilateral or UN agencies.

One document suggests the inclusion of more peer panellists and advisors from developing countries, as this could potentially benefit all concerned.

Familiarity with UNEG Norms and Standards and with UN evaluation functions are seen as important requirements in recruiting advisors for peer reviews.
One peer review panel comes to the conclusion that the established distinction between panel members and advisors should be reconsidered as it may be more practical to have panels in which some members undertake a larger amount of work than others.

7.4 Peer review framework

Most of the comments in the lessons learned documents on the peer review framework concern the first version of the framework (the EVALNET-UNEG framework developed in 2006) and many of these have been addressed in the second version of the framework developed and adopted by UNEG in 2011.

One example is the selection minimum UNEG norms and standards that a peer review should address as well as their grouping according to the assessment criteria of independence, credibility and utility.

One comment on the first version of the framework, which is mostly still valid for the present version, is on the issue of the effectiveness of the peer reviewed organization, and the contribution of the evaluation function to public knowledge about this effectiveness. This aspect, which includes the link between results-based management and evaluation, is seen to not be sufficiently covered in the framework.

Linked to the previous comment is the view that the peer review process could benefit from being more focused on strategic issues. One panel comes to the conclusion that there is need for prioritisation in the process and possibly for a simplification of the framework.

8. Analysis and discussion

8.1 Assurance and reliance

The peer review mechanism was conceived initially to deal with the fact that bilateral donors were not seen to be using the evaluations of multilateral agencies. The hypothesis was that this state of affairs was due to misgivings on the part of bilateral donor agencies about the quality and independence, among other things, of UN evaluations. As well, the peer review mechanism was seen as a way to develop evaluation capacity in UN agencies.

"Two key factors led to the introduction of Professional Peer Reviews of evaluation functions in multilateral agencies in 2004: a strong demand for multi-donor evaluations of UN organizations on the one hand, and the recognition of the need to harmonize evaluation practice due to the considerable variation across the UN System on the other. In view of this, the Evaluation Network of the Development Assistance Committee (DAC)/ Organization for Economic Cooperation and Development (OECD), jointly with the United Nations Evaluation Group (UNEG), introduced the Peer Review mechanism.

The rationale behind the Peer Review of the evaluation function of a UN organization is to establish the credibility of evaluation reports coming from the organization itself and thus potentially decreasing the need for external multi-donor evaluations of an agency or its evaluation office. In this way, the donor community can rely more on the multilateral organizations' own evaluations."[8]

The central idea of the peer review mechanism has been that it should provide assurance to donors on the quality of evaluation functions of UN agencies and the reliability of their evaluations thus resulting in less external evaluations of UN agencies by donors. Furthermore, peer reviews would improve the quality of the evaluation function which in turn would contribute to improving the effectiveness of the organisation.

---

[8] Peer review of the evaluation function of UNIDO (2010)

While the study findings suggest that peer reviews do contribute to improving the capacity and quality of evaluation functions, as well as generating greater trust in them on the part of bilateral donors, they do not indicate clearly whether increased reliance on, and use of, evaluations correlate with a reduction in bilateral evaluations of UN agencies that have been peer reviewed.[9] This analysis may suggest a gap in the initial logic of the peer review process, i.e. that increased assurance, reliance and use, would lead to a reduction in evaluations conducted by bilateral donors.

Interview respondents put forth a number of possible explanations for this possible gap, i.e. the peer review process does not engage the appropriate bilateral decision-makers, the credibility of peer reviews is insufficient, donors must respond directly to domestic political decision-making and accountability imperatives, and do so through their own evaluations.[10]

However, and although actions should be taken to address the first explanation and improve the second, a comparison to other peer review processes, primarily in the field of public audit, suggests that it may be unrealistic to rely solely on the current organisation of the peer review mechanism of UN agency evaluation functions to provide the kind, scope and level of assurance to bilateral donors that was initially intended and that may be needed to achieve a significant reduction in evaluations by donors.

8.2 Professionalisation

The peer review is a key component of the system of quality assurance and a mainstay of established professions. In this context, professional evaluation practise is relatively recent compared to other professional practises such as accounting and auditing. Evaluation does not yet have in place the legal grounding , the professional bodies and accreditations, the breadth, length and depth of practise, established professional norms and standards, the recognition of other disciplines and of the public at large, to name but a few conditions that apply to longer standing professions.

Although evaluation is undoubtedly on the road to professionalisation and significant progress is being made in developing individual, organisational and functional capacity for quality evaluation, the system of professional evaluation, generally and in the UN system, is still at a developmental stage.

---

[9] The study attempted to assess through other means whether there was a relationship between the quantity and nature of external evaluations of UN agencies conducted by bilateral donors and peer reviews of those agencies however methodological hurdles and resource constraints were such that the analysis was not pursued.

[10] Please see item 5.4.1 Use of UN evaluations by donors

From this perspective of evaluation as an emerging profession, and particularly with respect to assurance, other important aspects such as professional accreditation and public recognition, have not yet been achieved fully. In the context of professions peer review constitutes a necessary but insufficient part of a broader system of assurance for evaluation that can only be effective to the extent that other necessary parts are in place.

8.3 Lessons

Putting the study analysis in the context of professionalisation, two key lessons emerge:

The first overriding lesson is that the EVALNET-UNEG peer review mechanism "can't do it all" with respect to providing assurance on evaluation in the UN system. Other attributes of a professional evaluation system also need to be considered and developed and it is incumbent on UNEG to map out a strategy for the professionalisation of evaluation within its network, i.e. harmonisation, and for connecting with external professional evaluation networks, i.e. professionalisation.

The second lesson with respect to the current peer review process is that it should seek less to provide direct assurance, consistent with the specific role of peer review as part of a broader system of professional evaluation, and more to contribute to the professionalisation of its evaluation practise through peer assessment, exchange and support, internally and externally.

This re-balancing and re-clarification of the purpose of the peer review implies the adoption of an explicit agenda of professionalisation of evaluation as a basic frame of reference for adapting the current peer review content and process.

8.4 Peer review framework

Consistent with the preceding analysis, with the lessons learned from the reviews conducted to date and with the study findings, especially the importance of clarity of purpose of the peer review mechanism, the following considerations appear key for discussing and setting the future course of EVALNET-UNEG peer reviews:

The study found that, overall, the peer reviews adhered to the framework with the exception of engaging country partners systematically, due in part to resource constraints. The framework was found useful in part and more so for planning and organising the peer review process than for providing a basis for assessment. The study further found that the framework could be clearer about the nature of the exercise, i.e. that it is not an evaluation.

The peer review framework should be revisited so that it better reflects the role of peer review as part of an overall system of professional evaluation quality assurance and based, among other things, on peer involvement, i.e. a "professional model" of peer review.[11]

In particular the framework should make clear the purpose of the peer review so that it reflects unambiguously the appropriate balance between peer assessment, including exchange, learning and support, and provision of assurance. As indicated in the lessons learned from peer reviews to date, assessment criteria should be made more explicit. The references to UNEG Norms and Standards in the framework should be assessed for consistency and coherence with emerging professional standards in evaluation generally.

Peer reviews in professional practise areas such as public accounting and audit often consist of assessing the peer reviewed organisation's own quality control and assurance systems and practises, against a set of clearly established criteria. There is little reference in the UNEG Norms and Standards to questions of quality control and assurance. As such, they should also be reviewed in terms of their currency with up-to-date evaluation theory and professional practise and their utility and fit with, inter alia, professionalisation of evaluation.

8.5 Efficiency of the peer review process

Although the study findings show that the peer reviews conducted to date have been valuable, there is a general consensus among respondents that the process is too resource intensive, both in terms of direct costs, i.e. est. 90,000 to 200,000 USD, and of indirect costs, i.e. salaried staff time. As well, the length of the period of time over which the peer review takes place is considered by a number of respondents to have been too long and drawn out.

There have been improvements in the efficiency of the peer review process with, among other things, the introduction of a revised and more flexible framework as well as of a "reduced" or "light" option.

However, to the extent that the peer review approach is based on an assessment of the evaluation function's own quality control and assurance systems, that assessment criteria are clearer and more focussed and that peer assessment , exchange and learning are prioritised, there are opportunities for improving the efficiency of the process.

---

[11] The current framework has its origins in a peer review approach applied to States as set out in a 2002 document from the OECD Directorate for Legal Affairs. Although the current framework has evolved it still refers to a model of peer review more appropriately applied to States, i.e. OECD, NEPAD, WTO, than to professions.

The standard of evidence for a review is less than that of an evaluation (which is less than that of an audit) and is gauged to the level of assurance that the exercise intends to provide. With less of an emphasis on assurance and more on peer assessment, exchange, learning and support, peer reviews should require less time and resources to complete.

8.6 Equitable access to peer review

On timing of the peer reviews the lessons learned suggest that their value can be augmented if they are *"planned and scheduled to mesh with important decisions, reviews and/or relevant events within and beyond the agency being assessed."*

This said, the study found that generally it is the larger and more developed evaluation functions that have the ability, the resources, the profile and to some extent the confidence to request and volunteer for a peer review and to do so in a timely fashion.

The way in which the current peer review mechanism is set up would appear to put at a disadvantage smaller evaluation functions, which probably would most benefit from the capacity building and profile raising effects of a peer review.

As well, the relative significance to bilateral donors of different evaluation functions and their organisations affects their ability to obtain funding and resources for a peer review.

This apparent correlation, together with the risk that bilateral donor financing affect or be perceived to affect the impartiality or independence of the peer review process, suggests that a different funding mechanism should be developed. Such a mechanism should allow for equitable access to financing by all multilateral evaluation functions and should include a contribution of the peer reviewed entity, commensurate with its capability. A common pool of funds could for example be constituted by a regular annual percentage contribution of an organisation's evaluation budget.

Equitable access to peer reviews across the UN evaluation system would also allow for them to be made compulsory and conducted with regular periodicity, e.g. every 3 to 5 years.

The composition of the panel and the selection of panel members should also be reviewed. A smaller number of panel members, e.g. two, would reduce costs and likely increase the efficiency of the peer review process. Panel members could be selected more for their professional knowledge, skills and credentials than for their familiarity with the UN system (which may or may not be a good thing).

Funding from a common pool would do away with the apparent convention that the financing donor is part of the panel, a potential source of partiality and bias, and has a predominant say in the selection of the advisor.

8.7 Involving decision-makers

A number of respondents indicated that peer reviews did not necessarily or always "reach" the appropriate decision-makers in donor organisations, e.g. in multilateral departments. The study was not able to assess the extent to which there was systematic communication of peer reviews by EVALNET members to their respective multilateral departments and relevant colleagues however there appears to be some gaps in the process.

According to OECD DCD-DAC, EVALNET has developed a joint approach to fill the information gap on *"the institutional performance and development effectiveness of multilaterals. (...) This joint approach combines reliance on and strengthening of existing evaluation systems, including through peer reviews, institutional reviews by the Multilateral Organization Performance Assessment Network (MOPAN) and a new meta-evaluation/meta-synthesis approach."*

The study however did not find clear and explicit coordination mechanisms between these various initiatives and suggests that these be developed or strengthened to support their complementarity and the coherence of the accountability and assurance system.

Achieving this coordination and coherence will require however a common conceptual framework which, among other things, addresses in cogent fashion linkages between different levels of analysis, e.g. project, programme, functional, organisational, etc. and different functions, e.g. review, evaluation, audit, performance reporting, etc.

9. Conclusions

The study undertook to assess whether (or not) the reviews have made a difference in the use made of evaluations undertaken by UN evaluation departments by different stakeholders, in particular by donors; and whether these reviews have had a positive impact on the evaluation departments of the UN agencies which were subjected to such reviews. As well the study aimed to provide insights on the usefulness and feasibility of the Peer Review approach as outlined in the UNEG Framework for Professional Peer Reviews.

On the first question, the study was not able to conclude clearly whether, on balance, peer reviews have increased donor reliance on evaluations of peer reviewed UN evaluation departments to the point of reducing markedly their own evaluations of UN agencies as a direct result of peer reviews.

The study does however conclude that peer reviews have had a positive effect on donor and stakeholder perceptions of the value and credibility of evaluations undertaken by peer reviewed evaluation departments, as well as the accountability of their organisations. From these two preceding conclusions the study further concludes that peer reviews constitute a necessary but insufficient building block for increased and full reliance by donors and stakeholders.

On the second question, the study concludes that, generally, peer reviews have had a direct positive impact on reviewed evaluation functions as well as on their organisations, at both levels of governance and of management. The study further concludes that peer reviews have a positive and strengthening effect on the community of evaluators that spans bilateral donors, UN agencies and the evaluation profession.

On the third point, the study concludes that the peer review approach is both useful and feasible, based on the peer reviews carried out to date and in the progressively more flexible and adaptable nature of the framework iterations.

Building on these specific conclusions, and taking into account the evolution of peer reviews since their inception, the study concludes that, overall, peer reviews contribute positively to quality evaluation and through it to improved organisational performance.

However, the study also concludes that, overall, the positive effects of the peer reviews could be achieved in a more economical and efficient manner.

10.   Recommendations

10.1   Overall

The study leaves little doubt as to the value of the peer review mechanism for evaluation functions and their organisations both for bilateral donors and for UN agencies, and for the broader community of evaluation and evaluation stakeholders.

This said, the study makes the case for considering the peer review mechanism as part of a broader system of professional evaluation that is still in its developmental stage. The implication of this perspective is that the peer review process alone cannot and should not shoulder the primary or sole responsibility for providing assurance on evaluation quality.

Other components of the system of professional evaluation also need to be further developed, some of which are less amenable to the influence of the EVALNET-UNEG group than others.

What the study does recommend is that the initial raison d'être of the peer review mechanism, i.e. the provision of assurance to donors to foster greater reliance, be de-emphasised and given its appropriate remit as part of the broader system of professional evaluation and assurance.

It further recommends that the capacity building value of the peer review process be emphasised from a professionalisation perspective and with particular attention to the dimensions of peer assessment, exchange, learning and support.

10.2   Specific

The specific recommendations that are listed are summarised and referenced by number from the section on analysis and discussion:

8.3
UNEG should map out a strategy for the professionalization of evaluation within its network and for connecting with external professional evaluation networks.

The peer review should seek less to provide direct assurance, consistent with the specific role of peer review as part of a broader system of professional evaluation, and more to contribute to the professionalisation of its evaluation practise through peer assessment, exchange and support, consistent with the findings of the study.

This re-balancing and re-clarification of the purpose of the peer review implies the adoption of an explicit agenda of professionalisation of evaluation as a basic frame of reference for adapting the current peer review content and process.

8.4

The peer review framework should be revisited so that it better reflects the role of peer review as part of an overall system of professional evaluation quality assurance and based, among other things, on peer involvement, i.e. a "professional model" of peer review.[12]

In particular the framework should make clear the purpose of the peer review so that it reflects unambiguously the appropriate balance between peer assessment, including exchange, learning and support, and provision of assurance.

As indicated in the lessons learned from peer reviews to date, assessment criteria should be made more explicit.

The references to UNEG Norms and Standards in the framework should be assessed for consistency and coherence with emerging professional standards in evaluation generally.

There is little reference in the UNEG Norms and Standards to questions of quality control and assurance. As such, they should also be reviewed in terms of their currency with up-to-date evaluation theory and professional practise and their utility and fit with, inter alia, a professional model of peer review.

8.6

(…) a different funding mechanism should be developed. Such a mechanism should allow for equitable access to financing by all multilateral evaluation functions and should include a contribution of the peer reviewed entity, commensurate with its capability.

A common pool of funds could for example be constituted by a regular annual percentage contribution of an organisation's evaluation budget.

Equitable access to peer reviews across the UN evaluation system would also allow for them to be made compulsory and conducted with regular periodicity, e.g. every 3 to 5 years.

The composition of the panel and the selection of panel members should also be reviewed. A smaller number of panel members, e.g. two, would reduce costs and likely increase the efficiency of the peer review process. Panel members could be selected more for their professional knowledge, skills and credentials than for their familiarity with the UN system (which may or may not be a good thing).

---

[12] The current framework has its origins in a peer review approach applied to States as set out in a 2002 document from the OECD Directorate for Legal Affairs. Although the current framework has evolved it still refers to a model of peer review more appropriately applied to States, i.e. OECD, NEPAD, WTO, than to professions.

Funding from a common pool would do away with the apparent convention that the financing donor is part of the panel, a potential source of partiality and bias, and may have a predominant say in the selection of the advisor.

11. Annexes

11.1   Adherence to framework

Review of the evaluation function of:	UNDP, UNICEF, WFP, UNIDO, GEF, OIOS, UN-Habitat, UNEP, FAO (Summary of results)			
Head of Evaluation Unit:				
Date of report:	2005-2012			
Peer review panel:	Between 3 and 8 panel members			
Advisors:	1 or 2 advisors			
	Yes	No	Partly	Remarks
Peer review process				
1. Were the criteria for the selection of peer review panel members followed?	5	1	3	Some reports explicitly refer to selection criteria of panel members and advisors, others do not
2A. Did the peer review follow the suggested phases for the peer review process:				
a. Preparatory phase with formulation of ToR, development of normative framework and self-assessment of evaluation function	2		4	
b. Fact-finding phase with document review, preparation of a "factual report", first visit of the peer review panel, possible field visits, drafting of draft peer review report	5		1	
c. Peer exchange phase with a second visit to discuss based on draft report and finalization of report	3	1	1	Unclear for one of the reports
Or: 2B. Did the peer review follow the suggested process for a "reduced" process:				
a. Preparatory phase with formulation of ToR, development of reduced normative framework and self-assessment of evaluation function	3			
b. Integrated fact-finding and peer exchange phase with document review and one visit, starting with interviews and ending with peer exchange	3			
3. Appropriateness of peer review activities in light of the statements made in the report	9			
4. Did the peer review "seriously engage developing country partners" in the process of the review, the learning acquired and in the use of the results?	2	6	1	
Peer review content				
4. Did the peer review use the core criteria of independence, credibility and usefulness to assess the evaluation function?	9			
5. Did the peer review assess structural aspects of how the evaluation function operates in the organization?	9			
6. Did the peer review look at:				
a. policy issues?	9			
b. issues of planning and budgeting?	9			
c. coverage of the organization's achievements?	3	1	5	
7. Did the peer review look at:				
a. existing capacity and needs for capacity strengthening?	8		1	
b. the use of appropriate evaluation methodologies and the quality of evaluations undertaken?	9			
c. collaboration with local partners and stakeholders as well as harmonization and coordination with other external partners?	2	2	5	

11.2   Quality of peer review reports

Quality of peer review reports				
	Yes	No	Partly	Remarks
1. Is the report well-written and understandable?	6		3	
2. Is the report well structured (logical, clear and complete)?	5	1	3	
3. Are the purpose, objectives and scope of the peer review fully explained?	5		4	
4. Does the report contain a transparent description of the methodology used for the peer review, including limitations of that methodology?	6		3	
5. Are findings objectively reported based on the evidence?	7		2	
6. Do the conclusions present reasonable judgments based on findings and substantiated by evidence and provide insights pertinent to the purpose of the peer review?	5		4	
7. Are recommendations firmly based on evidence and conclusions?	7		2	
8. Are recommendations actionable and reflect an understanding of the evaluation function's organizational environment and potential constraints?	5		4	Question not easy to answer without having the necessary background on the organizational environment in all cases