



# Report of the UNEG Evaluation Practice Exchange (EPE) 2012 Seminar

**Dates:** 23-24 April 2012  
**Location:** Rome, Italy  
**Host agency:** FAO (with IFAD and WFP)

This report, prepared by the EPE Organising Committee, encompasses the discussions and outcomes from the UNEG EPE 2012.

## Introduction

1. The UNEG Evaluation Practice Exchange Seminar (EPE) 2012 was held at FAO Headquarters in Rome from 23-24 April 2012. The meeting was organised by the UNEG EPE Organising Committee (OC) co-Chaired by Emily Hampton Manley (OIOS) and Lori Bell (FAO). Members of the OC included Miguel Torralba (IFAD), Jock Paul (OCHA), Anna Guerraggio (UNEP), Jamie Watts (WFP), Shatho Nfila (ESCAP), Genny Bonomi and Matthew Fisher-Post (FAO).
2. Belen Sanz (UNEG Chair/ UN Women) opened the EPE by welcoming participants, both those in the room and those online. The EPE 2012 is the fifth EPE with 89 participants registered to physically participate in Rome. She thanked FAO for hosting the EPE and highlighted its importance as a unique opportunity for UNEG members to share practices. It is also an opportunity to facilitate learning, establish a community of practice and for the UN Evaluation family to come together. The EPE 2012 was designed to be both very dynamic and interactive with sessions to be held both in plenary and breakout/ small working sessions.
3. Bob Moore (FAO) welcomed all participants on behalf of the three host agencies (FAO, WFP and IFAD). He highlighted the increased importance of evaluation in FAO since 2010 when the independent evaluation office was created through the approval by its Governing Bodies of the evaluation charter. FAO has been an active participant in UNEG Task Forces (TF), which demonstrates their keen interest in learning and sharing information with UN colleagues.
4. Lori Bell (FAO) and Emily Hampton Manley (OIOS) welcomed participants on behalf of the UNEG EPE OC. They thanked the members of the TF for their valuable contributions and dedication to the event. Online participants were encouraged to send questions via the UNEG EPE email address. The three broad topics of the EPE – managing evaluations, results based management (RBM) and evaluation, and evaluation of complexity – were identified based on responses from UNEG members to a survey prepared by the TF, as well as from feedback from the AGM & EPE 2011. Discussion on the Draft Principles for the EPE prepared by the OC were also planned for the second day of the EPE and the outcomes were to be presented at the UNEG AGM (25-27 April 2012).
5. The plenary sessions of the EPE were webcast and an email address established for those web-participants to provide feedback (uneg-2012@fao.org). The recorded the sessions can be accessed at the following links:

### **23 April Sessions**

<http://193.43.36.192/20120423-UNEG-Evaluation-Practice-Exchange-2012-morning>

<http://193.43.36.192/20120423-UNEG-Evaluation-Practice-Exchange-2012-afternoon>

### **24 April Sessions**

<http://193.43.36.192/20120424-UNEG-Evaluation-Practice-Exchange-2012-morning>

<http://193.43.36.192/20120424-UNEG-Evaluation-Practice-Exchange-2012-afternoon>

6. The following session summaries cover the main points of the presentations and small group work. All PowerPoint presentations were posted onto the UNEG website at [http://www.uneval.org/calendarevents/eventdetail.jsp?event\\_id=308](http://www.uneval.org/calendarevents/eventdetail.jsp?event_id=308). UNEG members need to log on the site to access the material.

## 23 April 2012, morning session

### Plenary Session: Managing Evaluations

7. Introduction: Jock Paul (OCHA) chaired the session. He began introducing the presenters -Jamie Watts (WFP), Krishna Belbase (UNICEF), Adan Ruiz and Nurit Bodemann (MDG-Fund) and Juha Uitto (UNDP). “Managing Evaluations” as a topic attracted a lot of attention within UNEG – as a core set of professional activities for most UNEG members. The objective of the session was therefore to share experiences, innovations and challenges.
8. Innovations & Challenges in Managing Global Evaluations, Jamie Watts, WFP: In her presentation, Jamie Watts focused on WFP’s experience, innovations and challenges of managing strategic global thematic evaluations. These evaluations attempt to analyse experiences at a global or close to global scale and draw lessons for the organisation as a whole. The evaluations are global as they draw conclusions of WFP's work worldwide or provide a comprehensive overview of a range of issues over a 3-4 year period. The evaluations take approximately one year (start to finish) at an average of cost USD \$250,000. Final reports are presented to the WFP Executive Board.
9. Global evaluations focus more on learning than accountability and address issues that are evolutionary. A multi-country case approach is used and broader strategic questions asked, rather than the OECD DAC criteria, which focus on questions related to organisational capacity.
10. Challenges include focusing the evaluation, managing high quality qualitative data, combining global with national and regional or thematic analysis as well as cohesion among evaluation team members. WFP efforts to overcome these challenges include developing logic models, building capacity through the evaluation matrix, using evaluation teams that have a combination of evaluation and specialist expertise and holding post evaluation analysis meetings. These types of evaluations require more time for focusing and planning. Communication with the wide range of stakeholders is important because of the high learning potential these evaluations have.. WFP has developed a number of communication “products” to enhance communication and learning, including senior and multi-level debriefings, workshops to share and validate findings and recommendations, seminars and brown bag lunches etc. WFP’s Office of Evaluation has also been focusing on internal lesson learning through an informal “global evaluation focus group”, team leader workshops and peer-to-peer review and support.
11. Overall, global evaluations are particularly challenging and the opportunity to discuss and learn from other UNEG members was welcomed in particular on issues of resource limitations, team composition and management, and data management. They are also very relevant and of great interest

to Evaluation Office stakeholders (in particular senior management and the Executive Board, as well as to managers across the organisation).

12. Multi-Country Evaluations, Krishna Belbase, UNICEF: Krishna Belbase drew on experiences, issues and lessons from managing UNICEF's multi-country evaluations (MCIs). MCIs are thematic/programme evaluations with a country specific focus. A global term of reference is used but these are modified slightly for country case studies, to bring in country level realities. There are both global and national stakeholders and evaluation teams consist of both international and national expertise. The reports have a country specific management response and there is a good balance between learning and accountability. Examples of UNICEF MCIs include Tsunami Evaluation, Early Childhood Development, and Child Protection in Emergencies.
13. There are many challenges to consider when to managing this type of evaluations, from developing methods and design to managing the evaluation team to connecting with the variety of stakeholders. In terms of developing methods and design, a number of lessons have been learned – such as: conducting an evaluability study of four to five countries is impractical; programme design differences need to be tackled and multi sector evaluations are more complex; and there is limited scope for assessing impact. In terms of managing teams, UNICEF has experienced challenges in identifying and recruiting national consultants; keeping the team motivated; supervising teams at distance; and maintaining independence and credibility specifically at the country level. In terms of connecting, challenges have included connecting with governments; ensuring stakeholder participation; and maintaining independence.
14. This kind of approach encourages learning through cross-country comparison; development of imaginative designs; increased capacity development and greater opportunity to influence policies. Lessons highlighted included the need to plan ahead of time and allowing sufficient lead-time in each phase. The inception phase in particular is very central to the evaluation and the evaluation manager should participate in at least the first country visit. Theories of change and programme logic models should not be too complex. Evaluation managers should have some knowledge and experience of the subject matter and country context – not just evaluation methods. Evaluation managers should also have passion!
15. Joint Programme Evaluations, Adan Ruiz and Nurit Bodemann-Ostow, Millennium Development Goals Achievement Fund (MDG-F):. The MDG-F conduct four typologies of evaluations – joint programmes (JPs), focus country, thematic windows and global. Evaluations of JPs adopt a multi-sectoral approach and have a very complex structure in terms of outputs. They cost an average of USD 22,000 and are implemented over a three-month period. JPs' mid-term evaluations are funded by the MDG-F Secretariat while the same JPs in the countries fund final evaluations. Both mid-term and final evaluations assess the relevance of the JP, as well as its design and internal coherence; examine the JP's degree of efficiency in implementation; assess the extent to which the JP achieved its planned results; and measure the JP's contribution to the achievement of the MDG's, Paris Declaration Principles and UN reform. Final evaluations also identify best practices and lessons learned. Mid-term evaluations are oriented towards making practical recommendations for the programme implementation whilst the final evaluation is an assessment of how the JPs have been successful in achieving results.

16. In terms of use, the JPs are meant to improve collecting knowledge, institutional learning and accountability. There are four user groups – MDG-F Secretariat, donors, partner countries and UN partner agencies.
17. Conceptual challenges in the evaluation process include managing complex processes created by the multitude of partners involved; implementing quick evaluations and measuring results of JPs conducted over a short period (three to four years); and working in an environment where there is a limited evaluation culture. To overcome these challenges, the joint programme is the unit of analysis; participatory processes are implemented; a normative approach to evaluation designed and implementation was adopted; a centralised roster of consultants and procurement services was developed; and evaluations were managed from a central office in New York.
18. A standard methodological approach was adopted although the exercise was shortened. This approach included document reviews, interviews, direct observations and focus groups, and participatory workshops. A systemic approach to identifying consultants was needed to attract and retain talent and quality.
19. Some conclusions were highlighted: stakeholders find the evaluation a useful exercise to improve their evaluation capacities; mid-terms evaluations improved the management of JPs, strengthened joint work and accelerated their implementation pace; and evaluation findings have served as a basis for decision-making on the management of the Fund.
20. Assessment of UNDP's contribution to National Development Results, Juha Uitto, UNDP: The ADRs are independent evaluations of UNDP programmes at the national level to determine performance and strategic positioning. DAC criteria are used but have been slightly modified to match the UN perspective. The Evaluation Office of UNDP is moving towards a standardised approach for the ADR's but there is still some level of flexibility so it can be adapted as necessary. The objectives of an ADR include: providing substantive support to the Administrator's accountability function to the Executive Board; supporting greater UNDP accountability to national stakeholders and partners in the programme country; serving as quality assurance for UNDP interventions at the country level and contributing to learning at corporate, regional and country levels.
21. The Evaluation Office has undertaken 60 ADRs over the last ten years with the almost half being conducted since 2008, which has been possible due to the standardisation of the process. ADRs look at UNDP's contribution to national development results over one to two programme cycles. The evaluability assessment and the identification of stakeholders are important aspects of the process that relies heavily on a multiple-method approach. Each ADR takes between 6 months to one year and costs between USD 150,000 to 200,000. Interviews and focus groups are important data collection methods.
22. ADRs are an external process and help improve UNDP's performance in the country and to identify the organisations comparative strengths. They also foster an evaluation culture in the country. Some challenges of the ADRs include building national ownership; balancing accountability with learning; involving in-country stakeholders and ensuring utility. UNDP also had a poster on the same topic displayed during the coffee break.

## 23 April, morning session

### Managing evaluations small group work sessions and evaluation resource share fair

23. Following the plenary, participants were invited to participate in one of four group work sessions based on the presentations given during the plenary session. Additionally, there was an evaluation resource share fair where members (such as UNDP, UNICEF, FAO, IFAD, WFP) provided copies of reports and other publications for others to peruse and takeaway.

#### **WFP: Global/ thematic evaluations – Session moderated by Jamie Watts**

24. Twenty participants from CTBTO, ESCAP, FAO, GEF, ICAO, IFAD, IMO, IOM, OCHA, OHCHR, UN Women, UNODC, UNRWA, UNV and WIPO
25. Discussions focused on the challenges of scoping the evaluation and keeping the evaluation team focused once the evaluation focus has been decided.
26. Particular issues raised about “scope” related to the importance of engaging stakeholders in the scoping process, especially where the focus of the evaluation is not well defined and to managing stakeholder expectations.
27. Evaluation tools were seen as a key reference point for keeping the evaluation team focused during the evaluation phase. Some agencies also find it beneficial to have some EO staff participate in the evaluation teams to help retain focus. It was agreed that the careful selection of consultants was crucial and this led to discussions on the challenges in consultants who can manage these complex evaluations.
28. Issues to be taken forward to the UNEG Annual General Meeting (AGM) included identifying ways to improve access to qualified consultants and developing UNEG mechanisms for “scoping” information that may be of interest to others (e.g. UN Women “map” of joint gender programmes).

#### **UNICEF: Multi-country evaluations – Session moderated by Krishna Belbase**

29. Four participants from IFAD, OPCW, UNEP and WFP
30. A number of key learning points were identified during the group session.
31. Utility of the evaluation –
- Plan and spell out how and by whom the evaluation will be used.
  - Clear and targeted recommendations (feasibility and practicality).
  - Use of national reference groups for government response to evaluation recommendation and formulate lessons for similar contexts.
32. Country selection –

- Need to set criteria (data availability, capacity, representation, cost-efficiency).

### 33. Methodology –

- Global design that allows differences to respond to country context.
- Invest time and effort with first country.
- Select first country opportunistically.
- Keep broad parameters but adjust as the process moves forward.
- Do one country at a time.

34. The group proposed a webinar to further discuss the issue with a focus on relevance and lessons across agencies.

### **MDG-F Joint programme evaluations – Session moderated by Adan Ruiz and Nurit Bodemann-Ostow**

35. Ten participants participated from CTBTO, FAO, OCHA, UN Women, UNCDF, UNIDO and WFP

36. Questions were asked on the following themes: participatory approaches; independence; findings; management response; knowledge management; and evaluation recommendations. The group also discussed possible priority issues for a UNEG guidance document on joint evaluations.

37. The Fund's approach to ensuring a participatory approach to evaluation is to balance inclusiveness with simple and manageable implementation process. A number of partners are included in the process but that inclusiveness is limited to maximise the efficiency of the evaluation process.

38. To ensure the independence of the evaluation, the Secretariat (who commissions and manages the evaluations) does not have any decision-making power in the design or approval of the JPs, nor is it involved in programme management. Also, the evaluations are conducted by external evaluators who are not involved in the design or implementation of the JPs. The MDG-F does not coordinate directly with central evaluation units in its individual evaluations. Working directly with central evaluation units at HQ level would pose the risk of creating silos between agencies in the evaluation process. In its view, the responsibility of coordinating with central evaluation units lies within each agency's country office.

39. In terms of overarching findings, the MDG-F mid-term and final evaluations show that coordination among UN agencies in the MDG-F JPs has not been optimal and serious challenges to joint implementation remain. However, results vary by programme, and some have shown notable achievements in joint work. The Secretariat has noted some important achievements in strengthening coordination between UN agencies and national counterparts, and particularly between government ministries at central and local levels – while there is still a long way to go. Finally, the evaluations show that MDG-F JPs have provided effective and replicable governance models for joint work at the country level, through its National Steering Committee, Programme Management Committee, and

Programme Management Unit, which involve stakeholders from the donor and national governments, the UN RCO, and the various implementing agencies.

40. The MDG-F has developed a template (“improvement plans”) for the JPs’ management responses, which are submitted 15 days after the final evaluation report has been sent to the country. The Secretariat’s evaluation unit follows up with the countries on completion of the template. The Secretariat portfolio managers continuously monitor implementation of the improvement plans every six months. The country teams begin thinking about their improvement plans early in the process—already from the evaluator’s debriefing meeting during their in-country visit. This has helped in the finalization of the improvement plans after the evaluation. Management responses for mid-term evaluations obviously have different implications than for final evaluations. However, the Secretariat has incorporated a management response component into the JP final reports in order to encourage institutionalization of results.
41. The group then discussed priority issues that the UNEG Joint Evaluation Task Force should consider when developing the guidance note on joint evaluations. The group agreed that it would be a useful exercise to develop a typology of joint evaluations in the UN system, upon which guidelines could be developed to illustrate effective examples and models for implementation. In such an exercise it would be important to use a “one size does not fit all” approach.
42. In terms of knowledge management (KM), the MDG-F has a KM strategy involving nine different teams in various UN agencies, implementing separate KM plans for each of the Fund’s eight thematic windows, plus an additional plan for gender as a cross-cutting issue. The MDG-F M&E and KM specialist coordinates these nine plans. Each team conducts a number of activities, including workshops, web-dialogues, knowledge transfer agreements, and research papers to involve JP stakeholders as well as a wider audience to communicate key messages and findings from the Fund’s work. An important part of the KM products is based on the findings of the JP evaluations.
43. It was suggested that the MDG-F systematize evaluation recommendations by agency so that individual agency evaluation units can track implementation. However, the MDG-F purposely discourages assessment of JP performance and recommendations by agency in its evaluations, as this would reinforce a silo approach to programme implementation. In specific cases where there has a serious issue regarding the role of a particular agency in the JP, evaluations have pointed this out and made corresponding recommendations.

#### **UNDP: Country evaluations – Session moderated by Juha Uitto**

44. Seventeen participants from FAO, IFAD, ITC, OHCHR, UN-Habitat, UNDP, UNFPA, UNICEF, UNIDO, UNV, WFP and WHO
45. The group discussed the audiences for the country level evaluations including national governments, country offices and, in particular, Executive Boards. It also discussed the issue of ownership, the reference groups participating in the design of the evaluation and how the use of national institutions affects methodology. The group identified the need for more harmonised approaches and methods between agencies when conducting country evaluations. With regard to UNDAF and the role of UNEG, the group questioned the role of UNEG, in particular as concerns accountability considering

UNDAF are decentralised which in turn led to discussions on how to contribute to a strong evaluation culture.

46. Issues the group identified to take forward to the AGM or for discussion at the next EPE included:

- Harmonisation of approaches and methodologies of country level evaluations
- Exchange of experiences before conducting joint evaluations except UNDAF's

## 23 April 2012, afternoon session

### Managing evaluations small group work on sub topics related to specific evaluation steps

47. Participants met in small work groups to discuss good practice recommendations, panel evaluation method, evaluation work plans and managing consultants. These sessions were led by Miguel Torralba (IFAD), Seetharam Mukkavilli (IAEA), Maria Santamaria (WHO) and Anna Guerraggio (UNEP) respectively.

#### **IFAD: Good practice recommendations – Session moderated by Miguel Torralba**

48. Twenty-one participants from ESCAP, FAO, IFAD, IOM, MDG-Fund, OHCHR, UNFPA, UNHCR, UNICEF, UNRWA and WFP

49. A number of key learning points came out of the discussions in this small group. Participants agreed that achieving good recommendations requires planning and facilitation throughout the evaluation process. Transparency and participations also needs to be promoted. Moreover, good recommendations are key to enhancing the potential use of the evaluation. The importance of recommendations should not be overemphasized - benefits can also be derived from the process and much can be learnt from findings.

50. An issue identified for the AGM or the next EPE was the need to define more clearly the role of evaluation in monitoring the implementation of recommendations.

#### **Panel evaluation method – Session moderated by Seetharam Mukkavilli**

51. Five participants from FAO, IAEA, IFAD and WFP

52. Mr Mukkavilli facilitated the session and began by presenting the approach to “panel evaluations”. Panel evaluations follow the following steps – the evaluation officer prepares the evaluation TOR and identifies a panel of three to five experts that will be responsible for conducting the evaluation. These should be qualified panellists and balanced both in terms of gender and geography, and conflicts of interest should be avoided. Remuneration for the panellists is optional.

53. Panellists meet at Headquarters for a fixed number of days. Prior to the meeting, the evaluation prepares a meeting agenda and desk review material. The evaluation officer facilitates the panel and

provides guidance on the evaluation in terms of relevance, efficiency, effectiveness and sustainability. During the meeting, the expert panel conduct interviews and videoconferences. The panel is responsible for the final report.

54. The suggested issue to be taken forward for discussion at the AGM is that panel evaluations are suitable for research and development evaluations in technical organisations.

#### **Evaluation work plan – Session moderated by Maria Santamaria**

55. Twelve participants from GEF, ICAO, ITC, OHCHR, OPCW, UN Women, UN-Habitat, UNDP and WFP
56. The main question addressed among the group was – how do agencies select evaluation topics when the criteria are not specified in the evaluation policy. “Who” approves the plan, and their knowledge and capacity, is important. The criteria to apply depend on the kind of evaluation and those organisations with normative and operational mandates/ work have an even greater challenge. Criteria are not codified or, if they are, they are rather generic but the value of specific criteria is unclear.
57. The group identified a number of issues either for the AGM or the next EPE including: criteria must be established for selection of issues for evaluation, which should be categorised with specific indicators. A survey of what UNEG agencies are doing would be helpful.

#### **Consultant management – Session moderated by Anna Guerraggio**

58. Thirteen participants from FAO, IFAD, IMO, UN Women, UNCDF, UNEP, UNIDO, UNV, WFP and WIPO
59. The session was introduced by Anna Guerraggio (UNEP), who briefly illustrated the content of the “Enhancing evaluation quality through better consultant management” paper jointly prepared and presented by UNEP/UN WOMEN/UNICEF at the 3rd South African Conference on Monitoring and Evaluation (2011).
60. The group was asked to provide feedback on four issues: i) How to ensure overall evaluation quality through standard ToRs? Could internal peer review before ToRs are finalized be a useful tool in this respect? ii) What is the value added of the UNEG roster? How can it be strengthened? iii) How cost-effective is having “mandatory” inception reports? iv) Could evaluation reference groups increase the quality of the evaluation process and enhance learning and ownership of evaluation findings? v) How to ensure that the consultants team adequately involves stakeholders in the process? Does it only depend on its working conditions?
61. The discussion focused on questions ii) and iii). Group participants shared their experience and views on the use of inception reports as tools to better define the scope of evaluations, and the involvement of consultants in the development of such reports. While all the participants acknowledged the value of inception reports, different practices were noted. In particular, UNV asks the potential consultants’ team to develop an inception report before the assignment starts, and the quality of the report is used as a criterion for the team selection. Group participants also shared the challenges faced in hiring and managing consultants, who combine technical and evaluation expertise. Not all the participants were

informed of the existence of the UNEG roster, which is recognized as a useful tool for sharing information. Yet, the roster needs to be enhanced through regular and substantial inputs from UNEG members, and revision of some of its features (no “black list” for legal reasons).

## 23 April 2012, afternoon session

### Plenary session: Results Based Management (RBM) and evaluation

62. Introduction: Anna Guerraggio (UNEP) introduced the session. UNEG members showed a strong interest in the topic through the survey, although the OC were unsure how to broach it. A gradual change in the tasks demanded of evaluation functions within the UN has arisen due to a number of possible factors including: the introduction of RBM to the UN; measuring for results approach adopted due to the Millennium Declaration; and new public management and decentralization. She then presented the ladder of accountability (policy, programme, performance, process and accountability for probity and legality).
63. The introduction of RBM has brought greater attention to performance, and agencies have adopted strategic plans and RBM frameworks. Questions have been raised about the role of Evaluation Offices, which are often called to contribute to the definition of results and indicators, since they are independent and should therefore not evaluate something that they have set. However, expertise often lies within the Evaluation Office and it should be seen as a corporate role.
64. With the introduction of RBM, there has been a move upstream towards more strategic and results-level evaluations. At the same time, the decentralisation of the evaluation function in some organizations has brought about questions on how to ensure the quality of evaluations in the field and how the results/ findings from the field feed back into the big corporate picture.

#### **Feeding evaluation evidence on gender equality into decision-making and culture change – Session facilitator: Belen Sanz, UN Women. Panel members: Ashwani Muthoo (IFAD), Tullia Aiazzi (FAO) and Inga Sniukaite (UN Women)**

65. Introduction: Belen Sanz highlighted that evaluation of gender equality has been one area of attention within the UN system. Since 2005 there have been an increasing number of evaluations on these issues both in terms of performance (how organisations implement gender in their planning) and increased interest in capturing results on gender equality and empowerment etc. The panel members were asked to focus on three questions to address their experiences:
- Evaluation process: what was the evaluation approach deployed to evaluate gender equality?
  - Evaluation findings: how relevant have evaluation findings been against the underlying reasons for the decision to conduct such an evaluation?
  - Evaluation use: how have the results and recommendations of the evaluation fit into UN system-wide (and UNEG) mandates to integrate HR&GE in all areas of work, including evaluation?

**Feeding evaluation evidence on gender equity into decision-making and culture change: the Evaluation of FAO's work and role related to Gender and Development, Tullia Aiazzi, FAO**

66. Tullia Aiazzi addressed the three questions through a presentation on the evaluation of FAO's work and role related to gender and development. This was a corporate evaluation requested by the FAO Programme Committee conducted in 2010 and reported in early 2011.
67. The evaluation used the ECOSOC definition of gender mainstreaming and the notion of empowerment to analyse FAO's achievements in gender mainstreaming and women related activities. Nine per cent of the organisation's work was gender related which led the team to examine if the remaining 91% really was gender neutral. One of the team members also assessed the Evaluation Office's attention to gender mainstreaming.
68. FAO's action plans on gender and development had not been assessed since their inception in 2002. In 2009, the organisation developed a new strategic objective - "Gender equity in access to resources, goods, services and decision-making in rural areas". In parallel to the evaluation, FAO's Gender and Equality Unit commissioned a gender audit focusing on internal processes.
69. The evaluation found that progress on gender mainstreaming in FAO fell short of plans and expectations. A complete change in policy and accountability within the organisation on gender equality and mainstreaming was therefore required. The evaluation findings confirmed and strengthened those of the audit and reinforced each other's visibility resulting in a rapid update of the recommendations.
70. The main recommendations were that FAO should:
  - Re-state strongly at the most senior level, its full commitment to gender equality and mainstreaming gender as a corporate responsibility and not of one division only;
  - Establish an accountability mechanism at senior management level through RBM.
  - Gender should become a regular item on the agenda of senior management
  - FAO Governing Bodies should receive regular reports stemming from the accountability mechanism and recommend actions when compliance is not fully satisfactory.
71. The recommendations were accepted through the management response. Since 8th March 2012, FAO has a gender policy and senior managers have gender mainstreaming in their performance evaluation.

**Corporate-level Evaluation on Gender Equality and Women's Empowerment, Ashwani Muthoo, IFAD**

72. Ashwani Muthoo noted that this was the first corporate-level evaluation on gender by the independent evaluation office and was requested by the Executive Board. The findings were presented in December 2010.
73. The IFAD Gender Plan of action (2003-2006) outlines IFAD's corporate objectives for gender and gender mainstreaming across the organisation. In July 2008, the Programme Management Department

issued a framework for gender mainstreaming. In 2009, in view of IFAD's work on gender, it received the Global MDG3 torch committing it to doing something extra for promoting gender equality and empowering women.

74. The four objectives of evaluation were to: assess the relevance of IFAD's actions in promoting gender equality and women's empowerment; learn from the experiences and good practices of other development organizations; assess the results of IFAD-funded activities related to gender equality and women's empowerment in its country programmes and corporate processes; and generate a series of findings and recommendations that will assist IFAD Management and the Board in guiding the Fund's future activities in this area.
75. The evaluation had four mutually reinforcing building blocks. The first was an analysis of the main concepts and evolving approaches to gender equality and women's empowerment, and benchmarking. This entailed a literature review of IFAD documentation and what was available to understand the theory of change. Benchmarking helped IFAD to learn good practices of other organisations and see how they could be applicable to their work. The second was an assessment of IFAD's work to promote gender equality and women's empowerment at the strategic level from senior management. The third was a meta-evaluation of IFAD's past performance, collection from the field of evidence on the performance and perspectives of partners in borrowing countries, and review of recent country strategies and operations. The fourth was an assessment of selected IFAD corporate business processes (HR management, organisation architecture, quality assurance in design of projects and programmes, knowledge management etc).
76. The evaluation was organised in four main phases – inception, deskwork, visits to 10 IFAD-funded projects in five countries and writing of the evaluation report (including preparation of the management response).
77. A number of the evaluation's conclusions were given, including:
  - Gender equality and women's empowerment is essential for sustainable agriculture and rural development.
  - IFAD has played a leadership role and developed a comparative advantage, and is better than its peers at integrating gender issues throughout the programme and not just at the design stage.
  - Achievements on the ground are moderately satisfactory for two of IFAD's three strategic objectives (promoting women's access to assets and enhancing their roles in institutions to allow greater participation in decision-making). Achievements for the third strategic objective, improving women's own well-being, were seen as moderately unsatisfactory.
  - IFAD's strategic guidance was fragmented with no corporate policy on gender equality and mainstreaming. There was no common understanding across the organisation on terminology.
  - The strategic guidance is limited only to operational areas and does not cover corporate business processes. Corporate business processes are critical for better gender results, but they are mostly weak.

- Results from past IFAD-financed operations are on the whole moderately satisfactory
- Recent COSOPs and operations reveal improved design and performance
- The Executive Board has neither requested for reports on the performance nor ensured the allocation of necessary resources
- There is a gap between strategy and action: Are we walking the talk?

78. Recommendations coming from the evaluation included: IFAD management should, through a participatory process, develop evidence and results-based corporate policy on gender equality and women's empowerment, covering both operational and corporate business processes; there is a need for greater investment and staff training to build a common understanding on the theory of change and across the terminologies; management should get more involved in monitoring and reporting progress on gender equality and women's empowerment; IFAD management take a comprehensive review of the Fund's gender architecture; and the evaluation office should assess gender equality and women's empowerment in its evaluations.

79. Since the evaluation, the IFAD Executive Board approved the first corporate policy on Gender in April 2012. The policy covers both operational work and implementation strategy including gender and diversity balance in IFAD and includes a results and accountability framework (at least at the divisional level). A mid-term review is planned for 2014.

**Promoting accountability and the use of evaluation for gender equality results: in UN Women and UN system, Inga Sniukaite, UN Women**

80. Inga Sniukaite focused her presentation on – how UN Women promotes the use of evaluation results internally and UN Women's contribution to the system-wide coordination of evaluation?

81. UN Women's mandate is to assist member states progress more effectively on achieving gender equality. It has three work clusters – normative (support to member states decisions and General Assembly resolutions on gender equality), operational (global, regional country programmes on gender equality at the country level) and leading UN coordination (promoting accountability for gender equality results). In line with the coordination element of the mandate, the Evaluation Office evaluation strategy is to coordinate gender equality evaluation in the UN system through: UNEG; joint evaluation initiatives and building evidence based knowledge on gender empowerment and women's empowerment (GE/WE).

82. As UN Women's work is gender focused, evaluations are by default gender equality and human rights responsive. Evaluations specifically assess the extent to which the programme/policy evaluated is guided by organisational and system-wide objectives on HRGE; incorporate these approaches in the evaluation process; and understand evaluation as a political change process.

83. UN Women conducts corporate and decentralised evaluations and they are an integral part of the programming cycle. The evaluation office stresses the use of evaluation findings and managers are encouraged to use these findings when developing their programmes. Overall, the organisation is very clear how evaluations are used.

84. The Joint Evaluation of Joint Gender Programmes (JGP) in the UN system assessed to what extent and in what ways:
- Collaborating in a JGP enhances the GE/WE results achieved by the participating UN agencies and their partners.
  - JGPs are contributing to governments meeting their commitments to the Beijing Platform for Action and fulfilling their obligations towards women's and girl's human rights;
  - JGPs are contributing to improved gender equality mainstreaming and women's empowerment in other UN programmes and efforts at country level.
85. Two of UN Women's initiatives to promote better accountability for results in the UN were presented. The Joint Gender Programmes (JGPs) Portfolio Analysis is the most complete database in the UN on JGP with 113 entries. The Gender Equality Evaluation Portal is an initiative that systematizes information for meta-evaluations of gender equality results and knowledge platform. It contains evaluations explicitly assessing gender equality and/ or gender mainstreaming as a specific outcome of a project/ programme or as an institutional goal or strategy. Data is still to be evaluated and is therefore work in progress.
86. To conclude, a number of questions were posed on key concerns for gender responsive RBM practice regarding the use of evaluations.
- How and to what extent evaluation findings and lessons on gender equality issues are integrated into new programming cycles?
  - How and to what extent evaluations account for the implementation of UN mandate as regards Gender Equality?
  - How and to what extent UN keep track of progress on Gender Equality within its own system?
87. Belen Sanz drew three points from the presentations. First, there is a need in gender mainstreaming evaluations to define what is meant by gender equality. Second, the evaluations had to look at corporate level business processes and identify case studies that looked at results on the ground. Third, evaluators have to balance learning versus accountability issues. The evaluations presented demonstrated a lack of accountability and the fact the evaluations recommended the development of an evaluation policy is a way to strengthening this element.
88. A small group work on evaluation of gender mainstreaming followed in the afternoon session.

## **23 April 2012, afternoon session**

### **RBM and Evaluation small group work and poster session**

89. Participants were invited to participate in one of two groups for small group work and a poster session. For small group work, Marco Segone (UNICEF) and Oscar Garcia (UNDP) moderated a session

on decentralised evaluation functions. Belen Sanz and Inga Sniukaite (UN Women), Tullia Aiazzi (FAO) and Ashwani Muthoo (IFAD) led a session on evaluating gender equality in corporate cultures.

90. The posters presented were about Evaluation Quality Assessments in two different agencies: WFP (by Jan Michiels, Cinzia Cruciani, and Federica Zelada) and UNFPA (Louis Charpentier)

### **Decentralised evaluation functions – Session moderated by Marco Segone (UNICEF) and Oscar Garcia (UNDP)**

91. Thirty three participants from FAO, IAEA, ICAO, IFAD, IMO, IOM, MDG-Fund, OCHA, UN-Habitat, UNCDF, UNDP, UNEP, UNFPA, UNHCR, UNICEF, UNIDO, UNRWA, UNV, WFP and WHO

92. Issues that came out of the discussions that the work group put forward for further discussion at the UNEG AGM were:

- The evaluation function should be compliant to the UNEG's standards of independence, credibility and utility
- The nature of the evaluation function should be relevant to the nature and mandate of the UN agency
- How to secure appropriate systems are in place to support the decentralized evaluation function, including funding?
- Should UN (and UNEG) get prepared to move towards using national evaluation systems?
- What is UNEG's role in strengthening the credibility of UNDAF evaluations?
- Is it possible to have UNEG Guidance on strengthening the relationship between centralised and decentralised evaluations?

### **Evaluating gender equality in corporate cultures - Session moderated by Belen Sanz and Inga Sniukaite (UN Women), Tullia Aiazzi (FAO) and Ashwani Muthoo (IFAD)**

93. Nineteen participants from CTBTO, FAO, IFAD, ITC, MDG-Fund, OHCHR, UN Women, UNDP, UNEG Secretariat, UNICEF and WTO

94. A number of priority issues were identified by organisations during the small group work:

WTO - Evaluate the implementation of recently approved gender policy and evaluation of projects, which have a strong gender dimension.

OHCHR - Incorporate a gender advisor to the team and establish a mechanism to peer review strategic documents.

IFAD - Designate a gender focal point with a real gender background and develop a capacity building programme for evaluation staff.

FAO - Integrate gender competency in every single evaluation, establish consultation mechanisms to key stakeholders and establish regional/ decentralised UNEG working groups.

95. A number of recommendations for further discussion either at the next EPE or AGM were identified including:

- Methodologies for gender responsive evaluations should be developed and strengthened.
- Awareness should be raised about accountability for gender equality results in agencies.
- Good practices on the inclusion of gender in evaluation should be synthesised.
- A simple scorecard/ benchmarks for good practice should be developed.
- A meta-analysis of gender equality results in the UN system should be conducted.
- Gender questions should be integrated into the peer review process.
- Gender standards should be developed and included in the UNEG Norms and Standards.
- Regional peer mechanisms should be established.
- Expertise should be sought from UN Women.
- UNEG should establish a help desk.

## 24 April 2012, morning session

### Plenary session: Evaluation of complexity

96. Jamie Watts (WFP) chaired the session. The topic was one of three identified from the UNEG member EPE survey. It is a topic that interests many but raises lots of questions, in particular on the UN's capacity to deal with this cutting edge topic. Given UNEG's limited knowledge in this area, an outside speaker, Ben Ramalingam, was invited to present. He is writing a book on complexity sciences and international aid and has significant experience on evaluation and development from both an academic and theory/practical approach.

97. Ben's presentation focused on what complexity is and why it matters and implications for evaluation. He began by highlighting a number of challenges<sup>1</sup> still faced by evaluators and the overarching

---

<sup>1</sup> Challenges highlighted include: the cookie-cutter approach posing challenges throughout the evaluation process; the lack of data and monitoring analysis leaves many studies open to questioning; the absence of realistic and measurable goals, baselines and targets means determining outcomes and impacts frequently require "leaps of faith"; consultants have their own pre conceptions that lead to weak understanding of the subject being evaluated and the organisations themselves; evaluation recommendations tend to be unrealistic and overloaded; ambiguity and dissent is common (especially when things go wrong); use is still highly uncertain, despite well-established good practices;

messages these challenges bring, namely, that there is real potential value of synthesising lessons from UNEG, OECD and ALNAP case studies and peer reviews of evaluation functions and that evaluation itself needs to do more double loop learning (i.e. evaluators need to ask whether they are doing things right and are doing the right thing).

98. Public administration tools and techniques tend to incorporate assumptions (on systems, change, behaviours and relationships) not suited to the challenges that are being faced today. Since organisations are wedded to these assumptions it is hard to move away from them. Such approaches work well in certain situations i.e. “where there is a straightforward task to perform, a stable context and operating environment, identical, duplicate products and compliant, predictable and reliable parts”. The reality is that the world is dominated by complex systems – global economy, climate and environment, energy systems, information etc. – and that much of what is to be evaluated is not easy to define, not stable, has unforeseen consequences, socially complex etc.
99. So what alternatives are there to linear systems? Complex adaptive systems adopt a managerial approach to complexity and distinguish between simple (following a recipe), complicated (rocket to the moon) and complex (raising a child) assumptions. Complex systems research seeks to understand systems, networks, behaviours and dynamics particularly in four important areas: emergent systems, unpredictable change, adaptive and evolutionary behaviours and diverse informal networks.
100. Complexity has implications for evaluation at three levels:
  - Implications for evaluation models and assumptions:
    - o Need to recognize the differences between the closed predictable systems to open interactive ones, and where these apply in development and evaluation.
    - o Need to understand dynamics of change to develop better theories of change.
    - o Need to think about how change happens, incentives, mental models, institutions, political change through adaptive evolutionary change through institutional analysis – game theory etc.
    - o Network analysis: from atomised actors with formal relations to informal, social cliques and groups
  - Implications for evaluation management – Evaluation is often seen as a rational technical information-generating process however the reality is much messier. Evaluation management needs to consider complexity of evaluations themselves with many interconnected parts and dimensions (no two evaluations are the same).
  - Implications for the purpose of evaluations – Evaluation is traditionally seen as standing in contrast to creative dynamism often seen in core operational work. However, the purpose and position

---

budgets are growing but still not always matching ambitions; evaluation is threatened or undervalued by existing organisational cultures and selective use of evaluations; there is a lack of strategic leadership vision about evaluation, where it's positioned in an organisation and what it is for; demand far outstrips capacity, so evaluation departments are overburdened.

of evaluation need to be rethought. Evaluators also need to think about relationships, responsibilities and relevance.

101. Complexity-oriented evaluations:

- Develop new measures and monitoring mechanisms as goals emerge and evolve
- Provide feedback, gender learning, support direction or affirm changes in direction
- Aim to produce context-specific understandings that inform on-going innovation
- Support hunger for innovation

102. Ben made a number of conclusions to his presentation. First, that there is progress in evaluation but there definitely needs to be a move towards a double loop learning approach. Second, that the way evaluators think about the world and evaluations needs to change. Complexity and complex systems hold some answers. Third, there are a range of implications for evaluations, from methods, to management, to purpose and philosophy, which can be seen as a spectrum for engaging with complexity and need to challenge assumptions.

103. The presentation was followed by a short question and answer session where members of the audience asked questions about engaging rights holders in evaluation and making the evaluations more responsive; how to apply complexity to development programmes versus emergency response; how to establish boundaries; how to address complexity in the absence of monitoring and how to simplify complexity.

104. Ben indicated that there is a real issue about rights based approaches. There are some programmes where asking the community what they want out of an evaluation would be really important but there are few opportunities where this has been done. Evaluations do not necessarily look at downward accountability and so evaluators need to think about methods and approaches, and questions that are being asked. There is a false economy of asking evaluators to address shortfalls in spending for monitoring peace. Boundaries do need to be selected even if they already have a number of pre-assumptions. By challenging boundaries it is possible to get a clearer depiction of how the world works, but choices still need to be made. RCT's are only applicable to about 10% of what evaluators do and rigour is not a function of RCT's alone. Quantitative tools can be used to explore complexity.

105. The session was followed by presentations by UNEG members that focused on how some agencies are trying to apply these approaches in their work.

**Joint Evaluation of Refugees in Protracted Situations: How WFP and UNHCR handle complexity, Marian Reed, WFP**

106. Marian Reed presented the WFP and UNHCR Joint Evaluation of Refugees in Protracted Situations: WFP and UNHCR with a particular focus on how complexity was managed. She began by introducing the processes followed including development of standard TOR, QA system, and generic working logic based on internal/ external literature, and the application of a mixed method impact approach. The evaluation is focused on refugee camps and on key assumptions. A stakeholder

analysis was conducted to understand the different players and data collection tools developed after the inception visit.

107. The old model is used to show support to refugees over the past decade and is very much focused on “doing the right thing” to meet immediate needs. The new model accepts new complexity and asks “are we doing the right thing”? It is based on literature, policy and practice developed over the last decade. The evaluation looks at livelihoods beyond relief; recognition of a wider range of stakeholders; the possibility of self-reliance as a refugee; changing population profile as a refugee and inputs of two UN agencies working towards the same goal. However, the theory of change of cause and effect was maintained which did not capture well the dynamic element.
108. Evaluation managers designed a complex intervention logic based on WFP and UNHCR’s policies and joint memorandum of understanding. Developing it collectively was a challenge, and many of the assumptions proved to be context specific and not generic. The intervention logic helped the team to identify the assumptions to be tested. It was analysed in four time periods and the focus was on food assistance. The ultimate long-term impact was ‘durable solutions’ – repatriation, resettlement and local integration. There was a major “break” in the logic when it came to linking directly the food assistance to the long-term impacts showing a break in discourse between the two agencies aiming for the durable solution. The intervention logic, therefore, did not well capture the political, cultural and mind-set approach to change that a ‘complexity’ approach might. It did help to identify unintended effects and opportunities.
109. Marian then outlined the how the evaluation was managed including the roles of the evaluation managers and team. WFP led with joint decision-making and communication and the focus was very much on the camps and the interactions of refugees with host population. It was difficult to analyse outcomes related to education, health, and livelihoods without taking into account a broader range of inputs/activities undertaken by agencies in these sectors. Findings were focused on primary sources (mainly refugees and NGO’s, host population etc.) and the focus was on outcomes and impacts of points of congruence between the two agencies. A common process was used for evaluation quality assurance across the evaluations.
110. Areas that worked well included: using plausible cause and effect to better understand immediate outcomes; identifying feedback loops; comparing different contexts within a country and between countries and identifying common themes that emerged which challenged traditional approaches and assumptions.
111. The linear model was limited and a complexity lens may have led to more case studies of individuals or groups, better understanding why durable solutions are not considered options, except resettlement in a third country (less emphasis on studying cause and effect). The complexity approach could emphasize topics that were only handled as factors explaining the results and these tools might better explore WFP’s institutional and cultural assumptions regarding food assistance.

**Incorporating Complex Systems Analysis into Theory of Change Impact Evaluation: Neeraj Negi, GEF**

112. Neeraj Negi’s presentation focused on incorporating complex systems analysis into theory of change impact evaluation. More specifically he shared two examples of how complex systems theory

helped GEF in its work. The rationale for using complex systems theory was to better understand the phenomena that interventions seek to affect and to tailor tools and questions that help assess specific complex dimensions of the evaluandum.

113. The evaluative questions asked included: What are the key elements and chains of causality that will lead to GEBs (impacts)? What progress has been made in putting in place and linking those key elements? To what extent have GEBs been achieved? What are the pathways and factors affecting further progress to GEBs? What have been GEF's role and contributions? There were a number of challenges in responding to the questions due to time lags, nonlinearity and nested scales. In some areas conventional tools work fine but are constraining in others. Complex theory helps expand the number of options available for the evaluation in helping understand systems better.
114. GEF used complex systems to develop a theory of change to identify and describe critical elements, key linkages and interactions, causality paths and mechanisms and expected impacts. GEF also used these tools in social network analysis to understand the roles and interactions of agencies and GEF's role vis-à-vis others.
115. There are three elements to GEF's work – investment, foundational and demonstration. Linkages between areas are not necessarily linear, although there is a clear movement from foundational to investment elements. Complex systems allow GEF Evaluation Office to identify projects that have a number of different components, which affect different elements and different levels all simultaneously.
116. GEF applied governance of complex social- ecological systems to see if it is playing its role and how it can be tracked. They conducted survey of key actors asking about their relationships with other actors (funder-donor based or collaborating). Responses were mapped out to see what kinds of relationships exist and what GEF was contributing. The mapping showed that removing GEF from the mix weakens links between actors and highlighted services that GEF are providing in that particular area.

**Modeling Systems Dynamics: How useful is it for evaluating “Industrial Upgrading” projects?, Peter Lowe, UNIDO**

117. Peter Loewe and Sebastian Derwisch (PhD student/consultant) presented modelling system dynamics in UNIDO's work through a research project that is currently being tested. It is applied in a particular complex setting of industrial upgrading (in this particular example of leather in Ethiopia) although there are no conclusive results as yet.
118. Industrial upgrading is a complex environment because of the extent of the dimensional linkages. UNIDO reconstructed the intervention logic of the projects. The interventions are systemic in that it is not possible to have only one channel so it is complex from the implementation end. They used conventional tools and GEF's concept of impact drivers that when it comes to high level impact one should look more down the line parameters that could influence impact (e.g. look at impact on poverty). They reached boundaries of tools because type of setting requires a more complex tool so experimented with a model.

119. UNIDO are experimenting with the systems model to overcome a too simplistic understanding of industrial upgrading. To understand the interplay between export/ import, local market; competition between formal and informal projects and the trade-off between short and medium term poverty impact. UNIDO also wanted to demonstrate the need for multidimensional interventions and proper timing of interventions and questioned whether or not lessons can be enshrined in an interactive learning tool.
120. The tool is not meant to be purely analytical tool but to rather a communication tool between the evaluation office and implementers. The simulation tool has seven input variables (upgrading interventions), 12 output variables and built in external factors (to build scenarios). The input variables look at upgrading of production equipment, skills development, access to credit, strengthening of national quality infrastructure, logistics and customs infrastructure, "buy local" campaign and promotion of labour standards. Output variables included equipment, skills, productivity, costs, quality etc.
121. The model structure is based on the three elements required for production: equipment, labour and skills. They also result in the attractiveness of the products and subsequently to demand. Inventory coverage is a parameter defined by production and demand and regulates both supply and demand, and price. Prices are also influenced by costs (both fixed and variable) and both effect the profitability of production and desired investment, which brings the model back to full cycle. Some of the interventions that were worked with were equipment and skills upgrading, infrastructure support, logistics and customs support and NQS support. These can be modified in both times, intensity and length.
122. In terms of preliminary conclusions, it shows that: systems modelling is a practical way to cope with complexity; it enhances communication between evaluators and implementers on complex interventions; it is useful for generic "types" of interventions; it has potential to enshrine evaluation "lessons learned" in to the way the system is being configured.

**Evaluation of Complex UNDP Programmes and Themes in Complicated Contexts: Approaches and methods to enhance the validity of the evaluation of UNDP contribution to strengthening national capacities, Indran Naidoo, UNDP**

123. Indran Naidoo's presentation focused on the practical ways in which complexity is considered in the Evaluation Office with examples to how illustrate lessons learned have been fed back into evaluations. The Evaluation Office's evaluations are faced with the challenge of nested complexities the scope and criteria, evaluands, contexts, and diverse stakeholders and evaluation demands.
124. Validity criteria look at construct and content and meaningfulness as well as internal and external validity. Capacity development is defined "the process through which individuals, organizations, and societies obtain, strengthen and maintain capacities to set and achieve their own development objectives over time". The implication is an endogenous process that is country owned, dynamic change process with back and forth movements, unpredictable and adaptive to varying circumstances. The results framework is not valid for theme or construct; too static to capture dynamic change process like capacity development; and not able to address systems, sub-systems, interactions and inter-dependencies and emergent patterns. Open systems theory is an alternative and more valid

approach that capture complexity of the construct. It also retains some of the key criteria of traditional results framework for performance assessment

125. Capacity development is a means and is tied to specific things. The Evaluation Office have studied the question of how to understand capacity for developing “national development strategies”. This has become an important part of UNDP upstream work because as a knowledge organisation having strong capacity should improve capacity and development effectiveness. However, the key challenge is that there is no unified definition of capacity.
126. The Evaluation Office designs considerations in the evaluation of UNDP contribution to strengthen national capacities. Elements that need to be looked at in terms of design in an endogenous system include: operating premises guiding the approach in systems theory; procedure for the open system evaluation; case study methodology; synthesis etc.
127. In terms of methodology, the evaluation should focus on self-reflection by national stakeholders on what happened and the successes and challenges. It is also important to highlight that it is not an evaluation of the government. When focusing on how the internal national systems interact with the UNDP system, only after nationals have described the internal process and systems is there an assessment of how UNDP engages the development process and assessment of quality of performance. The significant difference in evaluation with these criteria is that performance was based on evidence from self-reflection exercises and not from pre-defined measures of success.
128. The Evaluation Office developed propositions and sub-questions to guide the generation of evidence and to structure the inductive analysis. So propositions and sub-questions were structured around the generic business process for developing national development strategy and planning and responsibility of ministries and sub-national bodies. The question on focus in UNDP was structured around UNDP’s own Toolkit for Managing Change based on the open systems theory. Complex theory teaches the need to negotiate.
129. This type of evaluation depends extensively on country case studies, which use explanatory case study methodology and open loop learning and concepts of feedback and emergence. Multiple methods are used to support both accountability and learning.
130. In terms of lessons, the Evaluation Office’s approach is appreciated by UNDP but there are both time and cost implications. There must be ample time for team and national consultant discussion and understanding of the approach and for testing the validity of the data. There must also be expertise among the team members including evaluators and case study methodologist: systems experts and behavioural and cognitive scientist for advisory; and national/regional consultants with a good knowledge of UNDP.
131. Participants were then invited to pose questions to the panel participants. The first question focused on what should be done with the analysis and how it can contribute to the design of future project designs. The panellists indicated that the analysis should be presented to management and could be used to identify key questions and areas that the evaluations need to cover. One panellist indicated that analysis is a form of validation of evaluation methodology.

132. There was much interest in the system being tested by UNIDO and participants were interested to know who is using the system, how expensive it was to develop and how did they achieve management buy in. Peter Loewe clarified that since they are currently experimenting on a small scale, senior management are not yet aware of the tool. In terms of resources and time, both are limited and so it has not been a costly system to produce. The team are positive that the system will work, in particular as a communications tool between evaluators and project managers. It will be particularly effective in synthesising lessons learned for multiple projects.
133. To conclude the session, Ben presented some of his reflections. He highlighted that these ideas don't tell evaluators what to do but simply present new ways of thinking in terms of insights, ideas, way of moving beyond existing meta-models and they point to all kinds of mind-sets and attitudes. Evaluators need to face up to complexity (instead of navigating away from it) and get the questions right.
134. A more scientific approach can be taken to evaluating complexity but evaluators need to challenge assumptions and the way they think about things.

*We cannot solve problems by using the same kind of thinking we used when we created them.*

*– Albert Einstein.*

## 24 April 2012, afternoon session

### Discussion of the Draft Evaluation Practice Exchange definition and principles

#### Definition and principles

135. The OC 2012 presented a proposed definition and principles of the UNEG EPEs, developed to provide guidance for future OCs. Participants were invited to provide feedback on the draft principles, which would then be presented for discussion and approval at the AGM. Comments and questions included:
- The EPEs are a very useful opportunity to reach out and learn best practices but there should be more exchange throughout the year. The OC 2012 had considered having exchanges throughout the year and to have a more interactive forum on the webpage. However time was short, and it would have required a lot of inputs by the OC members. To continue the exchange process, Emily Hampton-Manley volunteered to organise the first brown bag lunch in NY and other members from different duty stations, such as Geneva, also volunteered to do likewise.
  - Could the EPEs be linked to the Norms and Standards?
  - There is a strong temptation for “specific show and tell” at the EPE, but it would be more beneficial to focus on lessons and challenges. The EPE OC should also strive for cross relevance among agencies.

- The EPE could be a way to create an evaluation culture within the UN.

### **Participation in EPE**

136. There was a lot of discussion on who is, and should be, invited to participate in EPE. There has been much discussion in the past and, until 2012, only UNEG member agency staff were able to participate. Reasons for this included:

- There are already lots of associations, meeting societies where these open discussions can take place.
- EPE is for exchange of experience among UN evaluators and encourages rank and open discussions that may not want to be open to outsiders.
- The capacity to organise a potential large-scale event.
- It is an internal UN networking opportunity.

137. The concern was raised that if the EPE was opened up to some groups, then others may also wish to also participate and so perhaps a way to limit external participation would be to establish criteria. There could be value added by inviting someone external who has a very specific expertise.

138. It was mostly agreed that the EPE is an opportunity to see talents within the UN, which larger conferences do not always allow and that if the EPE became too big of an event then it would lose its “beauty”. Participants therefore concluded that participation in the EPE should be kept to UNEG members only.

## **24 April 2012, afternoon session**

### **Key learning and messages from EPE 2012 for the AGM**

139. The EPE OC drew together the key learning and messages from the main sessions as well as the breakout groups, that were to be forward to the UNEG AGM. Specifically:

#### **Joint and Joined Up Evaluations**

- Should UNEG develop specific guidelines on joint evaluations?
- Could we develop a better understanding of the approaches used to undertake individual country evaluations and try to harmonize – particularly with respect to our approach to national stakeholders (steering and reference group arrangements).

#### **Gender equality and human rights**

- Strengthen and develop methodologies for gender responsive evaluations

- Raise awareness about accountability for gender equality results in agencies
- Synthesize good practices on the inclusion of gender in evaluation
- Develop a simple scorecard/benchmarks for good practices (connected to system wide processes)
- Conduct meta-analysis of gender equality results in UN system
- Integrate gender questions in peer review process
- Develop gender standards to be included in the UNEG norms and standards
- Establish regional peer mechanisms
- Seek UN Women expertise
- Add gender into the UN evaluation peer review process
- Develop a scorecard to see how agencies contribution to promotion of gender equality
- Strengthen methodologies on assessing gender aspects of the programme
- Can UNEG provide support in reviewing for smaller evaluation units

### **The evaluation function**

- The evaluation function should be compliant to the UNEG's standards of independence, credibility and utility
- The nature of the evaluation function should be relevant to the nature and mandate of the UN agency
- How to secure appropriate systems are in place to support the decentralized evaluation function, including funding?
- Should UN (and UNEG) get prepared to move towards using national evaluation systems?
- What's UNEG role in strengthening the credibility of UNDAF evaluations?
- Is it possible to have UNEG Guidance on strengthening the relationship between centralized and decentralized evaluations?
- Guidance needed on the decentralized evaluation function

### **Knowledge Management**

- Improve information about evaluation sharing processes (from many of the working groups)

- Information about UN agency evaluation work plans should be readily available (survey of what UN agencies are doing, post workplan on individual agency evaluation websites or on UNEG website?).
- Update the UNEG fact sheet (and underlying analysis about UN agency evaluation function)
- Find ways to improve information sharing about evaluation consultants
- Consider capacity development being added to the EPE programme

### **Wrap up**

140. Some final comments were made during the wrap up. Overall, participants were satisfied with the formats, feedback and materials. Participants appreciated the format change in 2012 which included break out groups and poster sessions in addition to the classic panel and plenary programme. However, the standard of materials was a bit mixed. Participants also indicated that they would have liked a longer EPE (three days instead of two). Also, it was felt that the second session of small group work at the end of day 1 should have been similarly introduced by plenary beforehand for warm up.

## Participants List

Initial	Surname	Organization
T	AIAZZI	FAO
K	AIDNELL	IFAD
N	ASUKAI	UNOG
K	ATANESYAN	IFAD
M	BARUGAHARE	UN-Habitat
S	BECH	UN-Habitat
K	BELBASE	UNICEF
L	BELL	FAO
K	BELL-GRECO	WFP
C	BERTHAUD	IFAD
A	BHOURASKAR	FAO
N	BODEMANN-OSTOW	UNDP
G	BONOMI	FAO
M	BRUNO	FAO
S	BURROWS	WFP
A	CAPELLO	UNODC
L	CHARPENTIER	UNFPA
C	CONAN	WFP
C	CRUCIANI	WFP
E	DANTE	ESCAP
M	DE GOYS	UNIDO
C	DE VIVANCO	FAO
J	DOBINGER	UNIDO
N	DOLABJIAN	ICAO
L	DUBREUIL	FAO
F	FELLONI	IFAD
P	FIRBAS	CTBTO
C	FRANZETTI	IOM
S	FRUEH	JIU
A	FYFE	UNCDF
O	GARCIA	UNDP
S	GREEN	OCHA
A	GUERAGGIO	UNEP
E	HAMPTON-MANLEY	OIOS
M	JIMENEZ PONT	ITC
C	KIRK	UNICEF
N	KUMAR NEGI	GEF

F	LAMBERT	UNFPA
A	LARMOYER	FAO
P	LOEWE	UNIDO
E	LORA	FAO
M	MAKSUDOVA	FAO
J	MICHIELS	WFP
F	MINELLI	OHCHR
B	MOORE	FAO
D	MOUAFO	UNV
S	MUKKAVILLI	IAEA
A	MUTHOO	IFAD
O	MYARD	ICAO
I	NAIDOO	UNDP
R	NEPAL	OPCW
J	PAUL	OCHA
C	PERCH	IFAD
M	PICCAROZZI	FAO
M	PRAYER GALLETTI	IFAD
D	PRIOUX DE BAUDIMONT	WFP
M	READ	WFP
M	RIISKJAER	ITC
A	ROSI	UNHCR
D	RUGG	OIOS
A	RUIZ VILLALBA	UNDP
F	RUPP-GEMBS	IOM
R	SAHOTA	UNV
M	SANTAMARIA	WHO
B	SANZ	UN Women
M	SEGONE	UNICEF
D	SHALLON	FAO
R	SMITH	WFP
I	SNIUKAITE	UN Women
M	SPILSBURY	UNEP
I	SUAREZ	UN Women
R	TERBECK	UNRWA
D	THAPA	WHO
T	TORDJMAN-NEBE	UNEG
M	TORRALBA	IFAD
J	UITTO	UNDP

A	VIAJAR	IMO
A	VIGGH	GEF
J	VILA	CTBTO
J	WATTS	WFP
J	WORRELL	OHCHR
F	ZELADA	WFP
J	ZHANG	IFAD