UNEG
United Nations Evaluation Group

Evaluation Practice Exchange and Professional Development Seminar

7-9 MAY 2018 • FAO HEADQUARTERS • ROME

FINAL REPORT

Food and Agriculture Organization of the United Nations

IFAD
Investing in rural people

WFP
World Food Programme

Pibor

Contents

1. Introduction

The 2018 Evaluation week took place in Rome, Italy, between 7th and 11th May and was organized with the support of the Rome-based UNEG member organizations: FAO, IFAD and WFP. An EPE Committee was created in September 2017, to lead the preparation and organization of the event. The EPE Committee was coordinated by FAO and included one representative from IFAD and WFP[1].

The UNEG week comprised the following events:

Mon 7 May	Tue 8 May	Wed 9 May	Thu 10 May	Fri 11 May
Professional Development Seminar (PDS)	Evaluation Practice Exchange (EPE)	Evaluation Practice Exchange (EPE)	Annual General Meeting (AGM)	Annual General Meeting (AGM)

The EPE total number of participants was a record 163 while the aggregate attendance of the three training of the PDS was of 132 people.

The documentation related to the 2018 EPE can be found on the UNEG website here:
http://www.unevaluation.org/2018_UNEG_EPE.

A post-EPE survey was conducted in May 2018, just after the EPE, to gather feedback from the participants, and provide input for the organizers of the next EPE in 2019. The survey was completed by 73 respondents. The most relevant results of this survey are included throughout this report, and a more comprehensive report on the survey results is available in Annex 3.

This report, prepared by the EPE Committee, encompasses the discussions and outcomes from the UNEG EPE 2018.

1 Members of the EPE Committee: Amélie Solal-Céligny (FAO), Genny Bonomi (FAO), Mikal Khan (FAO), Aurélie Larmoyer (FAO), Deborah McWhinney (WFP), Fumiko Nakai (IFAD)

2. Professional Development Seminar

The Professional Development Seminar (PDS) was composed of three trainings running in parallel for one full day. The seminar was delivered in some cases by a combination of UNEG members and external experts. The topics were defined following consultation with UNEG members through the online survey carried out in November 2017 by the EPE Committee. PDS participants were asked to register to only one topic prior to the event.

According to the post-EPE survey, over 70% of respondents believe the PDS should be replicated at the next UNEG week. Several respondents noted that the training is fully in line with the learning spirit of the UNEG week, and in some cases even more valuable than some EPE sessions. However a one-day session could not be expected to go into depth, and some follow up between the trainers and participants would be useful.

Real-World Impact Evaluation – Applying IE methods creatively

Trainers: Anna Henttinen (WFP), Jo Puri (GCF) and Bidisha Barooah (3ie)

The training on Real-World Impact Evaluation (IE) provided an opportunity to discuss about how to apply experimental and quasi-experimental IEs methods creatively. The training included presentations, facilitated discussion and an interactive practical exercise.

Some of the questions considered were: how can experimental and particularly quasi-experimental IEs be undertaken in situations where conditions are not perfect? Should they always be undertaken or should alternative methods be used? What does a good impact evaluation design look like - what components does it need to succeed in answering the questions?

After an introductory discussion and some definitions, the trainers provided examples of creative real-world designs using quasi-experimental methods. The afternoon consisted of an interactive Impact Evaluation Design workshop, which allowed participants to develop their creative IE designs for real-world contexts and decide which methods they would choose and apply.

The training invited to consider options for designing creative impact evaluations in difficult (i.e. real-world) contexts where data may not be available or the context may be shifting. The objective of the workshop was to introduce the audience to main impact evaluation techniques, and share how they have been applied creatively.

In the post-EPE survey, 27 participants provided feedback on this training, and over 80% said the content had a high degree of applicability to their work, and over 90% expressed a positive view on the modality of the training. While finding the overall content useful, some participants however expressed the need for more tailored training on impact evaluation, based on the diverse evaluation models of UN agencies. For instance, some noted that the training was targeted at a 'hiring manager' role rather than an evaluator or junior staff. Most respondents noted the importance of the subject and expressed the desire to explore it more in the future.

Data visualisation

Trainer: Andy Kirk (independent consultant)

The 'Data Visualisation and Infographic Design' workshop aimed to provide participants with an accessible and comprehensive introduction to data visualisation and infographic design. The focus of the training was to teach the craft of this discipline, helping participants to know what to think, when to think about and how to resolve all the analytical and design decisions involved in any data-driven challenge. Across the session there were references for some of the most common, contemporary technologies but the emphasis was on the underlying craft, regardless of the tools or skills.

Whether we realise it or not, we are all frequent consumers of visualisation and infographic designs so improving the sophistication of how one reads, interprets and evaluates the effectiveness of such displays is a key literacy.

There were four over-riding objectives:

1. To challenge existing thinking about creating and consuming visualisation works, helping to clarify convictions about the differences between good and bad visualisation design.
2. To enlighten with an appreciation of the wide range of analytical and design options, including chart types, features of interactivity, annotation, colour applications, and composition.
3. To equip participants with an efficient workflow giving them the confidence to make astute choices based on sound principles and practical guidelines.
4. To inspire participants to elevate their ambitions, by broadening their visual vocabulary and exposing them to the latest techniques and contemporary resources for developing their data visualisation capabilities.

The content was delivered through a blend of teaching, discussion, and group practice. The practical exercises varied in nature from evaluating work, conceiving ideas, and forensically assessing design choices.

In the post-EPE survey, 11 participants provided feedback, and they all expressed positive feedback on both the modality and the relevance and applicability of the training. Some respondents noted the need for more examples of linking data visualization tools to specific evaluation practices.

An introduction to the practice of Sense Maker: the use of micro-narratives for design, monitoring and evaluation of international development programs

Trainer: Steff Deprez (independent consultant)

Sensemaker is an innovative narrative-based research, evaluation and decision-making methodology designed to generate actionable insights and guide interventions in complex systems and processes. SenseMaker is based on the collection and indexing of micro-narratives. These anecdotes, experiences or stories are self-signified by the storytellers. This means that respondents assign meaning to their own stories (self-interpretation) immediately after they have shared their anecdotes, experiences or stories through a set of questions (signifiers) rather than an external intermediary interpreting the narratives (common in qualitative approaches). SenseMaker implies the collection of a large amount of stories (+300 up to thousands of stories) to gain multiple perspectives on the domain of interest. The signification (indexing) of the fragments allows for quantitative pattern analysis backed with explanatory narratives. A set of well-designed signifiers (compiled in a signification framework) assists in revealing, comparing and contrasting patterns that cannot be detected by reading (or textual analysis) of the micro-narratives. The SenseMaker® software for the collection and analysis of data is developed by Cognitive Edge[2].

This one-day seminar was an introduction to the SenseMaker methodology. It was designed for participants that wanted to understand the principles and practice of SenseMaker and how it can be used in evaluation of complex international development programs. Participants were introduced to the different steps of the SenseMaker process including examples and experiences from different SenseMaker projects.

In the post-EPE survey, 11 participants provided feedback, and 8 of these (73%) expressed a positive view on the applicability of the material to their work and on the modality. Some respondents expressed the need for more tailored training material in the future, noting for instance the positive example presented by IFAD in the use of Sense Maker.

2 www.cognitive-edge.com

3. Evaluation Practice Exchange

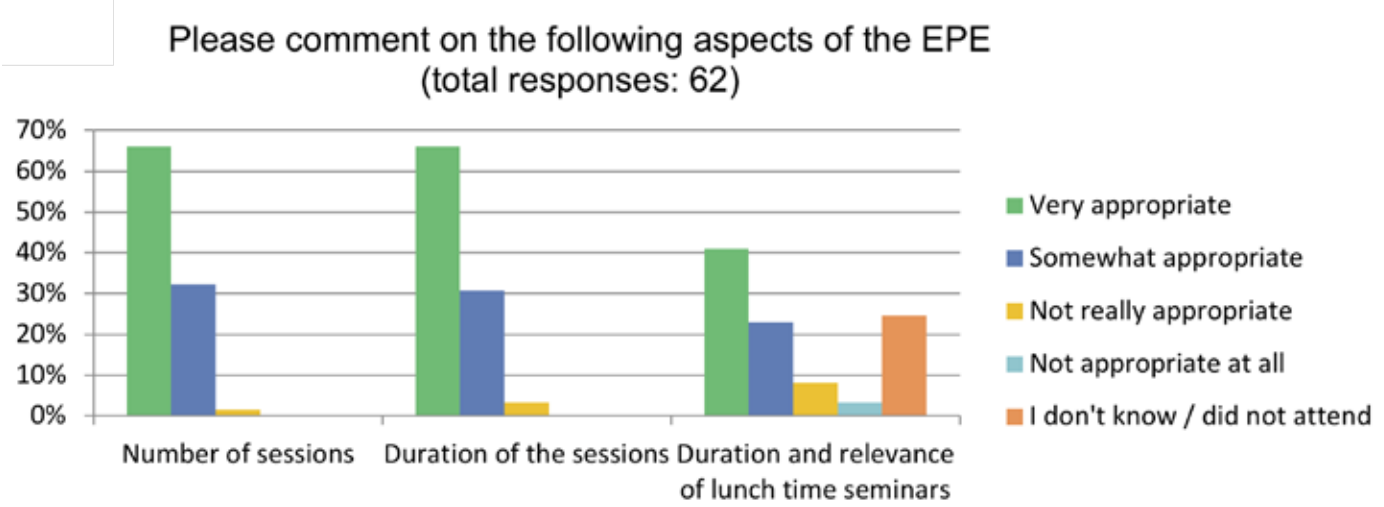
The 2018 EPE consisted of 18 sessions organized around three main themes: Managing Evaluation, Evaluation Principles and Evaluation Methods and various lunch-time seminars (see the EPE Agenda in Annex 2). The organization of each session, including the content and modality, was chosen by lead conveners, who were identified through an open call among UNEG members, coordinated by the EPE Committee. The lead conveners were assisted by collaborators. The EPE committee was responsible for the overall coordination and assisted the lead conveners with identifying collaborators and ensured that power point was avoided and that the content and modality were in line with the principles of usefulness and engagement.

According to the post-EPE survey, sessions with interactive modalities were often appreciated by participants who noted that the 'no-power point' rule should be kept. World cafés, Oxford-style debates, and role plays received very positive feedback from the participants. The move from traditional panel presentations to more interactive and engaging modalities was also welcomed. However, some participants pointed out that there is little value added in constantly sharing experiences over similar subjects, and would appreciate more focus on masterclasses.

Topics that interested participants the most were related to the following up of recommendations, data analysis and independence and the cost of evaluations. Interesting comment proposed to engage outsiders, specialist of the dedicated subjects, for some sessions in order to further raise the quality of the debates.

Quoting some of the survey respondents:
"It was great to move from the traditional panel presentations to more interactive sessions, there was so much cross pollination and I took a lot of very good ideas with me."
"Although it was well organized and innovative, it remains too insular and UN focused, and we need to draw more on what's going on outside the UN and involve non-UN people - it will higher up the quality".

Figure 1: Post-EPE feedback on number and duration of sessions



Innovations of the 2018 EPE

The 2018 EPE introduced a number of innovations which aimed at fostering greater learning and exchange among participants. This was done because of a demand expressed by participants of previous years' sessions, through a survey carried out by the UNEG Secretariat after the previous EPE in 2017 to collect the feedback from participants. In particular, five important innovations were introduced:

1. The Evaluation Practice Exchange was preceded by a one-day **Professional Development Seminar**, which provided a unique forum for UNEG members to get together face-to-face, to share and exchange experience, best practices and lessons and knowledge in evaluation, and to develop their expertise on specific subjects related to evaluation approaches and methodologies. Participants had the choice to attend one of the three training: 1) Approaches to Impact Evaluation 2) Data visualization tools 3) Sense Maker. A summary of the PDS was provided above.

2. In accordance with the results of the survey carried out by the UNEG secretariat following the 2017 Evaluation Week, the 2018 EPE focused on specific topics and issues encountered by evaluation professionals in their day-to-day work. Indeed **the EPE was about the bread and butter of evaluation officers**. It was aimed at helping evaluation professionals to better address some of the challenges we face as professional evaluators. The EPE session topics were selected based on the suggestions of UNEG members, and additional consultations among the EPE committee and UNEG secretariat.

The post-EPE survey found that the choice of the topics of the sessions was useful and instructive. On average 86% of participants found it either moderately or very useful.

3. A '**no power point**' rule was strictly enforced in order to promote more effective exchange and learning among colleagues. Lead conveners were encouraged to think creatively and use different modalities to organize their session in a participatory manner. Lead conveners were also encouraged to seek assistance from collaborators among the participants to help them organize and run the sessions.

According the post-EPE survey, over 90% of participants appreciated the no-power point rule and several comments across the survey acknowledged the enhanced engagement and knowledge sharing nature of the EPE. However some participants noted that greater engagement entailed less knowledge being delivered from conveners, and suggested a greater involvement of external participants to bring in new practices and ideas.

4. **Communication efforts were stepped up** with the use of videos before and after the EPE, as well as and graphic recording during the sessions. A short video (https://www.youtube.com/watch?v=jDA980O0fKY) was produced and shared with the participants a week before the EPE. In this video, all lead conveners presented a 30-second 'pitch' to present their sessions to encourage participants to attend. Another video was prepared upon conclusion of the EPE to document the proceedings and feedback with short interviews and images from the sessions.

   In addition to the video, the EPE introduced the use of **graphic recording** services[3] which were present in all sessions and produced visual representations of the issues discussed. The objective of this service was to contribute to the engagement of the participants and provide visual summaries for communication to UNEG members who didn't attend and to a broader audience. Finally, the opening and closing sessions of the EPE were webcast in order to allow all UNEG members to follow the proceedings[4].

According the post-EPE survey, participants and conveners overall appreciated the experience and would like the video format to be repeated for the next EPE. 86% of the lead conveners who completed the survey declared that the video was a good idea; 71% of the participants found it useful. Several respondents also noted that the video was good for a having a broad overview however the detailed session description was needed to have a full understanding of the content of the sessions.

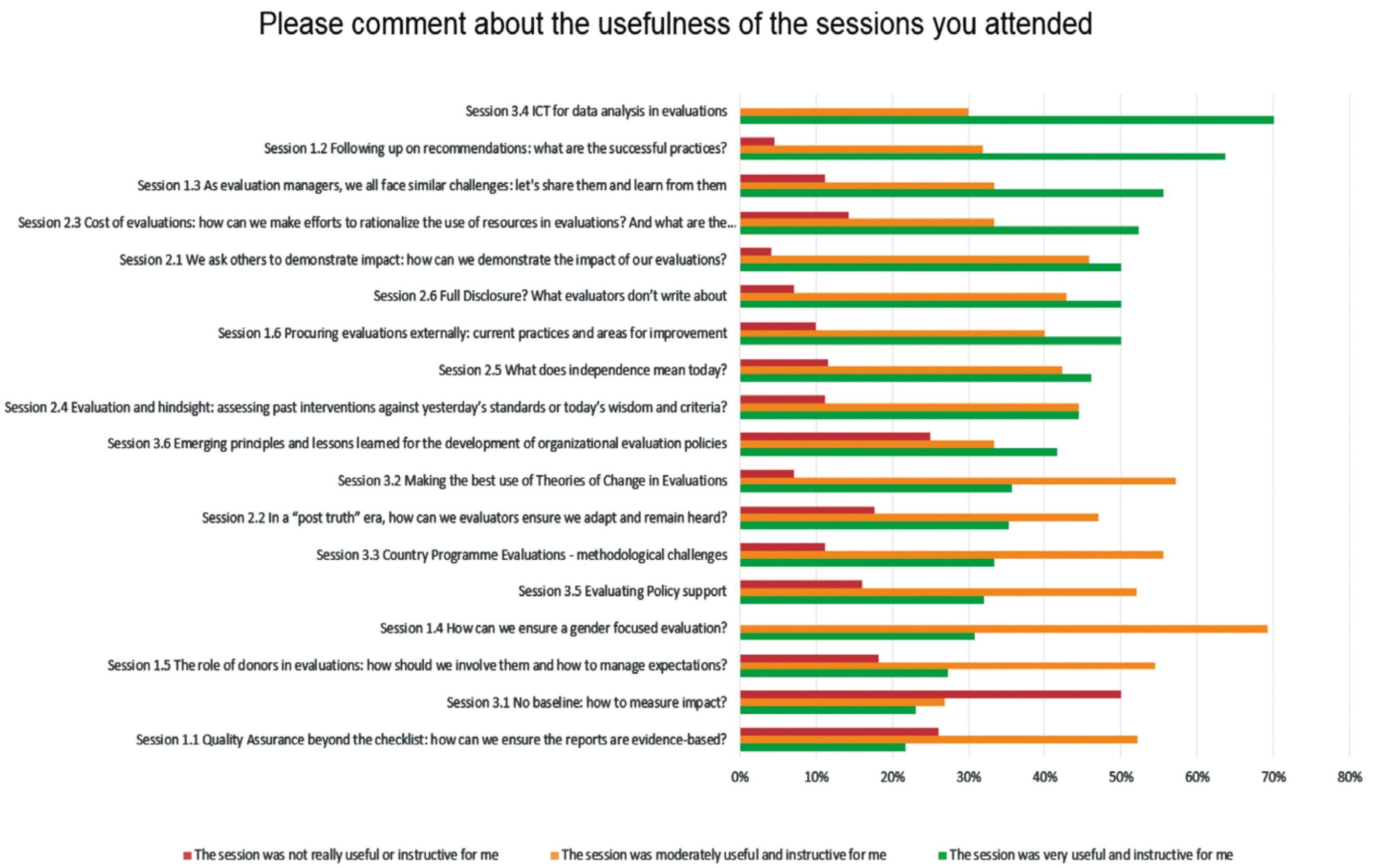
5. This year's EPE made use of the **Mentimeter**[5] web application to collect real time feedback from participants during the sessions using mobile phones and tablets on questions posed by the facilitator. Examples from the opening and closing sessions are shown below. The size of the words reflects the number of participants who typed it as a response to the open questions.

3 www.scribing.it. All graphic illustrations are available here: http://www.unevaluation.org/event/detail/490
4 Opening session: https://www.youtube.com/watch?v=3WWVtD8s33k
Closing session: https://www.youtube.com/watch?v=bafcb0KOMZ4
5 www.mentimeter.com

Figure 2: Mentimeter Opening session: In one word, what is an evaluator's most important skill? (106 responses)

Figure 3: Mentimeter Closing session: How would you describe the 2018 EPE in one word? (81 responses)

Figure 4: post-EPE feedback on usefulness of sessions



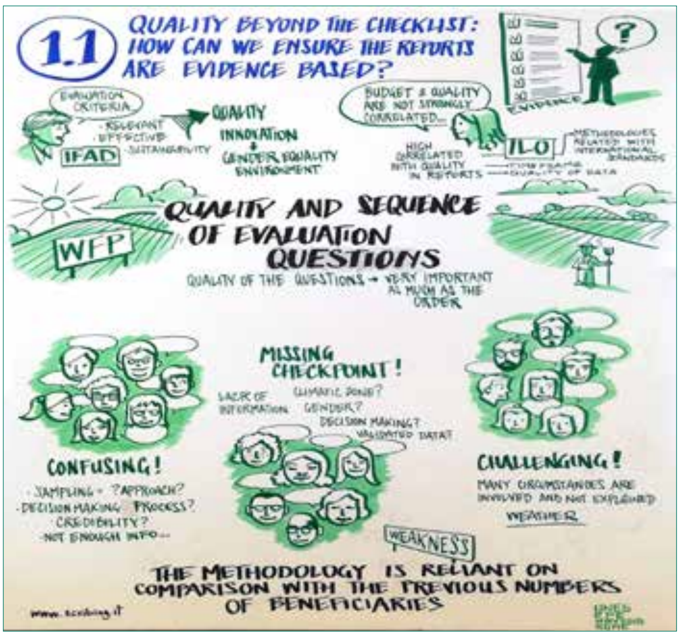
Theme 1 – Managing Evaluations

Theme 1 dealt with various aspect regarding to the management of an evaluation, from the recruitment of the team, to the dissemination of the report and the follow up of the recommendations.

UNCDF convened a **session (1.6)** to discuss issues regarding the "hiring of evaluators". Different modalities exist to procure evaluators such as long-term agreement, roster, individual contracts, or relying on firms. The session started by categorizing these different modalities and groups. It was observed that each modality has its pros and cons and there is no one-size-fits-all modality; this makes decisions context-specific and dependent on the organization and the evaluation. Different methods also apply to the selection process, ranging from selection matrix to reference checks. Debriefs to those who are not selected was discussed, as the question relating to briefing either with the firm or with individual consultants. Concerning the need to manage the contract during the entire evaluation cycle, timing is an uncertainty as it can change while the evaluation is running, which will impact the firm of the consultant. A key lesson is that it is important to build trust with the people that evaluators work with.

In **session 1.4** led by UNWOMEN, participants discussed ways for "designing gender-focused evaluations". They were asked to identify the five main challenges they face in trying to ensure gender-focused evaluations (impressions were captured using Mentimeter). Once the challenges were identified, participants were divided into groups to consider each one of them. A first key message resulting from the session was that evaluation managers have to begin with understanding what the specific gender issues in the project, the country, or the program are. Secondly, evaluators have to assess to what extent these issues are significant in a country or a project, especially for technical projects; it was remarked that these issues should not be forced anywhere. Indeed, gender-focused evaluations are very important but they may not always be possible in technical projects. It is determinant to understand the cultural context, the dynamics, the relations between men and women in a country and the impact on what one is evaluating. Sometimes, there can be underlying power mechanisms that are also very difficult to evaluate. Consequently, it was concluded that one needs to have quality information beyond traditional sex-aggregated data, which implies that there is a need for qualitative information regarding what are the gender issues in the evaluand. In that regard, a preliminary assessment is genuinely important.

Session 1.1 was convened by IFAD and dealt with "ensuring the quality of the reporting process" of the evaluation. UN organisations have different systems to ensure the quality of the reports, which was not defined as a problem. Organisations notably all have peer reviews process which are useful in ensuring the quality of the report. IFAD added some criteria to the usual ones; they added 10 notes to their specific mandate which allows them to look at criteria such as specific impact on rural poverty, gender or partnerships. Another discussion concerned the meaning of quality report: indeed, it was observed that one can have strict mechanisms for quality assurance, however, this does not necessarily mean that the evaluation will have an impact. Therefore, influential evaluations are not necessarily the same as high quality evaluations. Reflections thus concern what we want and how do we measure this.



In the same vein, **session 1.2** carried by WIPO considered "following up on recommendations" and different systems from several organisations were presented. Some organisations use quantitative approach, following tracking system number of recommendations that are accepted and implemented, and then they report the data through to their governing bodies. Other organisations rely on more qualitative systems that consist of dialogue and consultation both during and after the end of an evaluation. For instance, between four months and two years after the evaluation period there is a formal dialogue with the evaluands to discuss about the implementation of the recommendations, including their continued relevance, the status of implementation and any other issues. However, before thinking of following up on recommendations, evaluators need to make sure that they make actionable recommendations that are relevant, and that there are clear responsibilities. A crucial point to clarify is to identify who is the person in charge of following up, and who is in charge of implementing the recommendations, especially in case of final project evaluations – since after closure, the project staff will disperse.

Session 1.5 by UNESCO concerned the "role of donors in evaluations". It was observed that it is important to clarify the expectations of the donors regarding the evaluation and regarding how they are going to use it. Therefore, the evaluation managers must be clear from the beginning. Evaluators should bear in mind that there could be potential conflicts concerning what one should evaluate and the evaluation questions, as donors have different ideas from the evaluation manager – it is consequently determinant to discuss and negotiate if needed. However, donors are not a homogenous group as some may be interested in evaluation while some are not necessarily concerned. Some donors are also technically skilled and might want to influence the evaluator's design. Identifying the type of donor is a key stake – in case where the donor is not interested in evaluation, it was decided that there is no point to force him. In that same session, a discussion on MOPAN was also hold. Positive and negative feedbacks on MOPAN were shared by organisations. Negative remarks emphasised on the fact that MOPAN is a very long process, which limits its usefulness. Positive remarks concerned the fact that MOPAN was an opportunity to triangulate with other types of assessments from other organisations. The MOPAN process has recently been reviewed and will hopefully improve.

Eventually, **session 1.3** led by WFP concerned the "challenges that evaluation managers can encounter". The session was organized as a game, which consisted of a team competition choosing the best answer to a set of evaluation management challenges. These challenges ranged from anticipating the delivery of the evaluation report to managing fights within the evaluation team. Other topics concerned finding solution to the drop out of a team member during the fieldwork, or dealing with complains that the project team addressed to the director regarding the findings of the evaluation report. The conclusion was that challenges are very similar despite the differences in the organizations and that there is usually no right or wrong way to respond to these kind of challenges: this is part of the life of an evaluator. However, some important points can be kept in mind, the first one being to keep calm: it is just an evaluation! Secondly, a good evaluator needs to have a good understanding of methodologies, principles and the subject-matters, but also team management and conflict resolution skills as there is a high probability that there will be a need to manage some conflict at some point during an evaluation. Moreover, one needs to be flexible in order, for instance, to adjust the initial ToRs or methodology, based on the overall objectives of the evaluation. As the director of the FAO Office of Evaluation reminded us in the EPE opening session, one needs to be humble: we are just evaluation managers trying to do useful work.

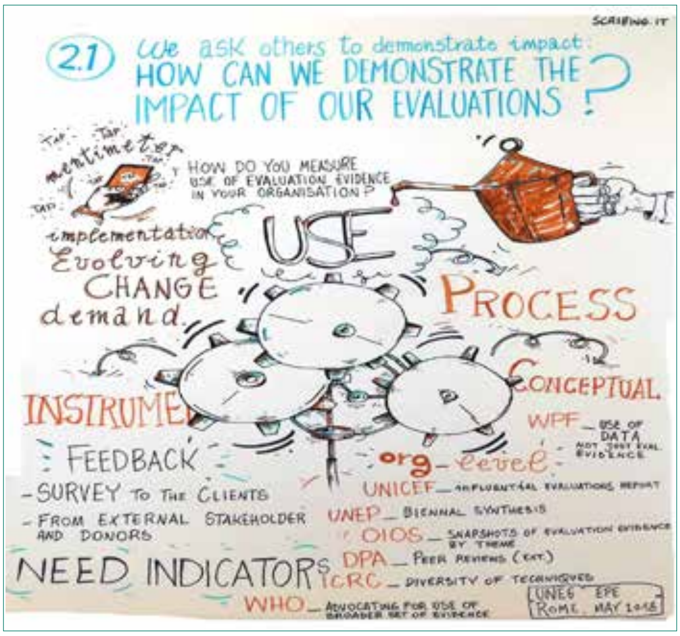
Theme 2 – Evaluation Principles

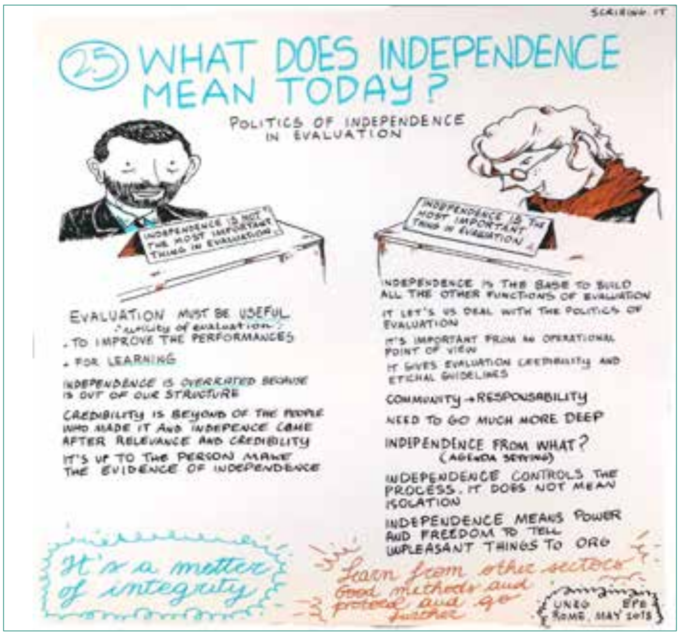
Theme 2 covered the topic of evaluation principles: sessions dealt with independence, post-truth era or full disclosure.

Session 2.4, convened by IFAD, was entitled "Evaluation and hindsight: assessing past interventions against yesterday's standards or today's wisdom and criteria". The session considered how evaluators assess past interventions against yesterday's standards or today's wisdom and criteria. Participants and conveners discussed different scenarios relating to changing project context, policies, and evaluation methodology change; before reflecting on ways to tackle these changes. One of the key messages is linked to the importance of analysing and understanding the context and the nature of the change. Indeed, one question for debate was about whether the change was formalised and documented, and to what extent it was formalised and documented. Conversely, in case of the absence of formalisation and documentation, the session considered to what extent it could be considered that this absence was reasonable. Moreover, the session emphasised the importance of being aware of the sensitivity around some of the issues of change, of the importance of the evaluability assessment of the data, of the capacity of implementers and program managers to adapt to the changes that happened, and, eventually, of the importance of assessing whether the change was reasonably predictable. That being said, it was noted that evaluators have to take into account the context of changes, which implies that a one size fits for all answer cannot be delivered. Another important point was about striking a balance between accountability and learning: it was noted that in some cases where sensitive issues may be raised, evaluators can focus on the learnings in reporting the evaluation findings.

Session 2.3 was entitled "cost of evaluations: how to rationalize the use of resources in evaluations". FAO convened this session which consisted in a fun roleplay with two scenarios. One of them consisted in a debate between an evaluation manager and a project manager in which the latter had received the preparation budget for an evaluation and was complaining that the evaluation budget was too high. The identified high cost items were notably the data collection exercise and the workshop, as well as the recurring use of international consultants. Different people were invited to intervene for the role of either evaluation manager or project manager. The main message was that costs are relative to the use, utilisation, and value of the evaluation exercise. Surely, there is an imperative to be cost effective for evaluation managers, however, it is the context that will drive the costs. For instance, participants recalled that, as project managers can regularly argue that they have already been generating data during the program and that consequently, the utility of collecting new one is very low, it was still up to the evaluation manager to judge of the quality and relevance of the available data and therefore, to estimate if new collection was needed. The quality of the evaluation manager will therefore rely on his understanding and designing of the evaluation in such a way that it takes into account the given context and the data available. In terms of use of national consultants, participants reflected on the equal association of the national consultants and emphasised the utility of national consultants to understand the context. The second scenario of this role-play staged an evaluation director and a chef de cabinet, the latter controversially arguing that he thought to get rid of the evaluation function as it was too costly, and as the Audit Office was becoming very similar to the Evaluation Office. The debate was genuinely animated and the question of whether the audit functions were getting closer and closer to the evaluation function was not entirely solved. Indeed, some people argued that the resemblance with evaluation should be taken seriously as the audit was not as much compliance-based as it used to be.

In the same vein, **session 2.1**, convened by WFP, also questioned the use of evaluation and was untitled "We ask others to demonstrate impact: how can we demonstrate the impact of our evaluations". Two development lines can be isolated; one was about the ability to measure the use of evaluation and was related to the kind of system or approach that can be used to observe how the evaluations are actually used. Different colleagues shared their experiences. To summarise, it was exposed that some of these approaches are formalised (such as the standard management response to reporting on the status of implementations, or annual reports asking the management to talk about the use of evaluations) while others are informal. In that regard, one participant mentioned an impact logbook where she would note down every time she observed something that, according to her, was influenced by the evaluation. The point was that the use of evaluation is broader that solely recommendations; it includes evaluation findings and remarks that are more analysis in evaluations than recommendations per se. Another important point touched upon obtaining the client feedback. It can be through formalised events: for example, external evaluators may be asking client feedback about how they have used or have not used evaluations in order to build up from answers. A second group conveyed another development line on the way the use of evaluation could be enhanced. A first response was logically related to the quality of evaluations but discussions also covered stakeholder engagement and the efforts that can be made to respond to the demands from the potential users on the kind of evaluation that are led.





Session 2.6 was entitled "Full disclosure? What evaluators don't write about" and was convened by FAO. Three main subtopics were discussed in the different groups. A first one dealt with fiduciary aspects and touched upon situations where evaluators encounter evaluation findings related to fiduciary issues but are not sure if they should write about it. A second topic concerned the project management leadership. It dealt with situations where evaluators find that the project leadership in effect affected the effectiveness of a program, and discussed different ways to behave in front of this situation. The third topic concerned political events such as elections or some of the issues a government can be sensitive about. Participants considered situations that may strike the balance between transparency and credibility. The session therefore stressed the importance of understanding the sensitive issues and focusing on the context while discussing example of leadership.

Moving on, **session 2.2** untitled "In a post truth era, how can we evaluators ensure we adapt and remain heard?" was convened by WIPO. Post-truth relates to alternative facts and so-called facts based on opinions, emotions, perceptions, or social media. Three axis were highlighted. A first one presented this world-wide phenomenon and discussed its implications for the evaluation design. The second axis refers to the implication of the post-truth context for the evaluation analysis, the conduct of evaluation and the implementation of the evaluation. The third one considered the implications for the evaluation communication. All topics are interrelated and stress the importance of being aware of the areas where sensitivity can come into play and of identifying the areas where some backlash can come from in order to build protections and pre-empt this backlash from the alternative paths. Notably, participants expressed the need to build protection into the evaluation designs from the beginning. Another argument concerned the need for a communication strategy that is participatory and engage stakeholder in the process. Attention should also be given to getting the buy-in from the people or the parties that would demand fair independent evaluations not influenced by emotions and alternative facts. Here again, the importance of the demand for credibility and evidence based on the facts was emphasized.

Eventually, **Session 2.5** "What does independence mean today?" convened by the GCF group reviewed the difference between independence of institution and professional independence. Independence in evaluation was notably highlighted as only one of the aspects of evaluation, among others key elements such as, notably, impartiality, credibility and utility. Some participants argued that independence should not be perceived as more important than those key elements, since building trust is the determinant factor. Independence seems to become critical only when it comes to exchange situations.

To summarise, the key messages from the sessions introduced above include i) the sensitivity to the context and the capacity to identify the context that causes the change, ii) the critical role of doing good evaluation design, notably for transparency, iii) the critical issue of the use and value of evaluation impact, that is, defining the purpose of the evaluation, iv) the balancing learning and accountability, v) the importance of stakeholder engagement in different evaluation processes. These are sort of education principles, norms and standards that sessions encountered all along.

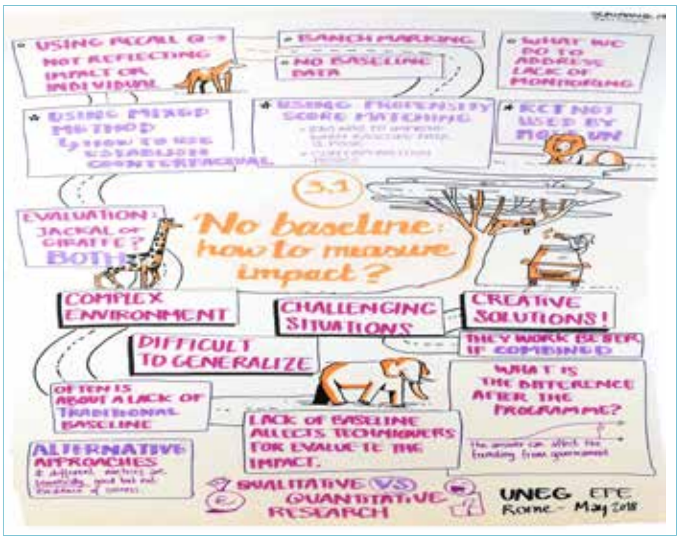
Theme 3 – Evaluation Methods

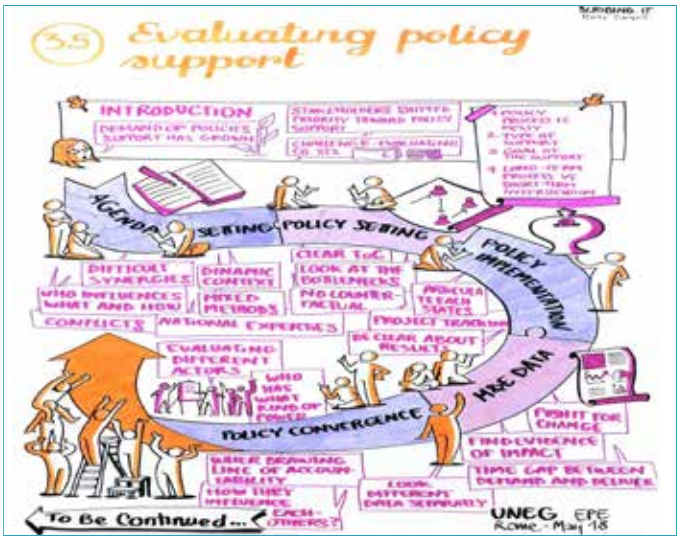
Theme 3 covered various evaluation methods such as impact assessment, the use of theories of change in evaluations, ICT solutions for data collection and methods for evaluating policy support.

It seems appropriate to begin with **session 3.6** which was untitled "Emerging principles and lessons learnt for the development of organisational evaluation policies" and was convened by GCF. Participants focused on the foundational concepts of evaluation policies which are credibility, utility and independence, using Mentimeter to provide feedbacks. Credibility was linked to issues of quality and the importance of the latter; the challenges and threats to the evaluation function within organisations; and issues of budget allocated to the function of evaluation. Regarding utility, the importance of participation, engagement, and relevance were emphasised. Concerning independence, the session discussed functional and behavioural independence, which must be considered as complementary aspects. Some of the takeaways relate to the need to develop a strategy to implement a policy and then monitor track and report on that implementation. There is also a need to have a policy that responds to the character and the nature of the organisation, and reflects its dynamic character and specificities. Moreover, it was observed that, if the relevance and purpose of evaluation were critical, maintaining independence was also critical and the efforts should be oriented towards balancing between the two. Eventually, participants also stressed the need to have management responses that are dynamic and not simply a pro forma exercise.

The **session 3.1** "No baseline, how to measure impact" faced some logistical issues and was therefore redesigned in its modalities by colleagues from FAO, UNIDO and OIOS. It was a safari style with groups split on different topics and it focused on the tools that can be used and the methods and approaches that can be taken when no baseline is available. One way is to use the theory of change, asking questions of 'how' and 'why', but not to focus on the numbers or the data. Instead, the session invited to be creative by asking other questions, using collaborative approaches and determining what to emphasise while still being able to evaluate the program or the project, and using establishing counterfactuals in creative ways as well as using tools like propensity score matching. The final takeaway point was that impact evaluation are not always necessary or even the best way as there may be alternatives that are as impactful and useful.

Session 3.2 "Making the best use of theories of change in evaluation" led by IFAD was divided into three groups. The first topic of the session concerned the value-added of using theories of change in evaluation. It notably relates to increasing the credibility of the evaluation, helping managing expectations, using ToC as a tool for sustainability, helping to establish an initial hypothesis and then being able to draft evaluation findings and establishing a causal link and preconditions for potential impact assessments. Participants then looked at the ways to engage stakeholders in the design and in the validation of that ToC. Some organisations do engage stakeholders regularly hence it was not something that had to be developed as a separate process, whereas others may have to design it and build it more formally. The validation of ToC with staples should be done quickly and carefully, taking into account levels of awareness about the tool and, in certain cases, as there may be political sensitivities involved, evaluators were invited to be aware of potential tensions amongst stakeholders. The third topic looked at the helpful tools that can be used to reconstruct ToC to support an evaluation. These included shifting the focus from outcomes and outputs to results and cause and effect relationships; moving from a log frame to a ToC approach and accompanying people so they understand what the implications are; brainstorming with the team; writing down and acknowledging assumptions; drawing a line; setting up a contribution ceiling for each step; using colour coding to demarcate pathways; and avoid being fixed on a ToC, taking into account the need to be iterative and revising as the evaluation process unfolds.





Moving on, **session 3.4** "ICT for data collection and analysis in evaluation" focused on the variety of tools and applications. The session was convened by UNITAR. It seems convenient to summarise here the pros and cons of these tools and applications. The highlighted pros were related to instruments such as the geospatial tools, considered powerful to quantify environmental challenges and to reconstruct counterfactual scenarios. Some felt that the use of GIS data increased the statistical vigour of the impact evaluation and helped to overcome issues related to baselines. Moreover, participants saluted the fact that these tools are often free, open sourced and easily accessible by many, and can be used for various quantitative analysis and complement evaluation methods – rather than replacing them. Some of the cons or challenges relate to the fact that they may need high degrees of computer power and technical skills; the fact that there can be uneven availability and accuracy of contextual variables; the fact that it can be difficult to answer the 'how' and 'why' questions; the fact that methods require really good M&E systems; and the fact that they can be less useful for quantitative analysis.

Session 3.5 "Evaluating policy support" unfolded by defining five stages of policy support and then broke participants into groups affiliated to the different stages. The latter refer to agenda setting and awareness-raising, policy spending or policy design, policy implementation, M&E and data, and policy convergence. Instead of going through each one, it seems convenient to mention general challenges about this policy support cycle. It was observed that it can be challenging to understand what is often and evolving policy context and to keep up with the actors that are involved as they regularly change. Understanding the evaluability of specific aspects of policy support can also be challenging. It is difficult to handle attribution and contribution as there is often a lack of clarity blurring the borders of what is the support to policy making. It is complicated to distinguish between outputs and outcomes as well as establishing M&E system within a fluid context and to measure impacts in this area that has a lot of greyness. Some of the solutions the groups came up with were related to the importance of doing an analysis of political economy, having a policy influence mapping at early stages with key players and key influencers, setting up monitory systems to understand what these key players are doing and what role they are playing in the agenda setting or policymaking, ensuring that each approach is tailored to the country's context, involving evaluators who are not only special to some policy support but who have a solid knowledge of the political economy of a given context, having a clear evaluation theory of change while being flexible to change approaches and assumptions if that policy context does change, and finally, to tailor the analysis to the type of policy support being evaluated.

In **session 3.3** "Country Programme Evaluation – methodological challenges" which was convened by WFP, groups were divided according to different challenges. The first one was related to balancing quantitative and qualitative methods and some of the challenges dealt with the validity or usefulness of data, addressing evidence gaps in a timely manner (i.e. how to catch up in a situation with uneven qualitative vs quantitative data, and vice versa). The identified solutions referred to having minimum requirements in order to establish criteria and standards linked to the amount and type of data that are to be collected; building monitoring frameworks; building capacity to increase quantitative data when the latter are missing; or increase engagement with stakeholder when qualitative information are missing. Another group discussed the theory of change and the challenges relating to the fragmentation of projects and programs when it comes to evaluate a country portfolio as it is sometimes difficult to have one theory of change if the latter was not already constructed; therefore, it can be necessary to reconstruct or adapt one to suit the needs of a country portfolio evaluation. A third group dealt with engaging stakeholders, stressing the importance of inclusive and participatory engagement despite time constraints. There needs to be an iterative approach to identify stakeholders and to include them in the process. Strategic communication with the country office was also considered as a key priority milestone. Participants also notified the interest of an interagency collaboration on field missions in order to understand which actor has been to a particular country not to trip over each members. A fourth group, concerned with availability of data, to as a starting-point an optimal scenario whereby evaluators would engage with government and have access to national statistics and perfect dataset, and can collect information from the country office on all their activities and results. More realistically though, evaluators would have to define proxy indicators that provide sufficient evidence for the evaluation to go forward, use data from other organisations, be an audit related date or date from IFRS, while also recognising the limitations of those datasets. In the same vein, participants noticed that evaluators can start with an evaluability assessment that gives an idea of where the gaps are. A fifth area is cross or interagency collaboration and evaluation. It was noted that there is often a lack of coordination, knowledge and sequencing of UNDAF evaluation on the one hand and country portfolio evaluation by various agencies on the other hand. It was also emphasised that there needs to be greater coordination starting at the central level, sharing work plans and genuinely engaging each other agencies about what plans are going to go ahead and how to better collaborate. It was suggested that decentralised evaluations in interest groups, as part of UNEG, could take this stuff where they look at the locus of the country and see how best to improve that kind of knowledge and information sharing. Eventually, the last group looked at scaling up as many organisations scale up to reach coverage norms. UNDP, which as experience in that topic highlighted the utility of sending pre-mission questionnaires to country offices to get information gathered ahead of time, reducing missions to field and being selective about the time they spend in the field, and developing regional rosters of evaluation consultants that organisations can pull in to help balance and contribute to that process.

4. Lessons learned for future UNEG weeks

The UNEG week 2019 will take place in Nairobi, Kenya, and will be organized by UNEP and UN Habitat. This section aims at providing useful information to the future EPE organizers, based on the feedback received from the 2018 EPE participants, during the event and through the online survey.

Preparation phase

The use of a survey of all UNEG members to select topics of interest was very valuable and was a key success factor for a useful EPE, and should be replicated. The survey was launched in November 2017, 6 months before the start date of the EPE 2018. This allowed for sufficient time to analyse responses, select topics and integrate them in the planning.

The creation of an EPE Committee in September 2017, 8 months before the start date of the EPE, with members from the three Rome-based agencies, was essential to ensure that the organization of the UNEG week had clear responsibilities and focal points in all the organizing agencies. The members of the committee had to plan their workload accordingly as the time invested for the UNEG week progressively increased approaching the event.

Modalities

Around 90% of the respondents to the post-EPE survey declared they would like the no power-point rule to be kept for the next EPE. Modalities such as World cafés, Oxford-style debates, and role plays received positive feedback from the participants. The move from traditional panel presentations to more interactive and engaging modalities was appreciated.

The selection of meeting rooms is important to promote participation and interaction. The post-EPE survey showed there was some dissatisfaction with those meeting room where seating and tables were fixed and could not be shifted to create small groups, or acoustics were not ideal, or there were no windows. In the case of the FAO building this was a limitation on which the EPE committee had no control.

Several participants also noted the need for more clarity ahead of the event on the expected roles of lead conveners and collaborators of the sessions.

Topics

The topics need to be carefully chosen based on needs and preferences of participants. These can vary broadly and the EPE should aim at addressing this diversity. In the 2018 EPE the topics that received the most positive feedback were: following up of recommendations, data analysis and independence and the cost of evaluations (for full details see figure 4 above in the summary of EPE).

Appendices – Session summaries

Session 1.1 – **Quality assurance beyond the checklist: how can we ensure the report are evidence based?**
Lead convener: **Oscar Garcia (IFAD)**
Collaborators: **Patricia Vidal (ILO), Diego Fernández (WFP)**

This session addressed the difficulties faced by Evaluation Managers to ensure good quality reports beyond the use of checklists. The experience of three United Nations agencies in establishing quality assurance standards and checklists for centralized and decentralized evaluations was shared. Afterwards, participants were provided with a case to reflect on ways to ensure final evaluation reports are making good use of evidence. The interactive session ended with a summary of lessons learned and areas for future research.

Modality

The conveners from IFAD, ILO, and WFP presented their institutions' evaluation structures and individual insights on factors impacting quality of evaluation reports. These presentations also included a description of each of the institutions' approach to ensure quality control in their evaluations: IFAD focused on quality control of project completion report validations, ILO on the governance structure supporting higher quality evaluations, and WFP on the design, structure and sequencing of evaluation questions.

Key messages

The mandate of **IFAD**'s Evaluation Office includes, among other things, the validation of the completion reports of all IFAD project, which are produced by the programme managers. Project Completion Report Validations (PCRVs) apply the usual evaluation criteria (relevance, effectiveness, efficiency and sustainability), but in addition they assess the projects against some criteria of specific interest to IFAD. This is an example of ensuring quality evaluation reports, beyond the standard OECD DAC criteria, and it provides more useful information. The criteria are:

- Rural poverty impact;
- Innovation in supporting new approaches to rural poverty reduction and scaling up;
- Partnership: IFAD, and the government;
- Quality of the Project Completion Report;
- Candor: was it an honest report or not (rating: 1-very bad, 6: very good. 4 and above are good).

In addition, there are facets influencing the quality of the report that can be isolated, notably: the rigorous use of evidence, the novelty, and an enhanced understanding enabling not just learning from current intervention, but for future interventions, enabling the conduct of an insightful synthesis.

ILO presented the structure of its evaluation function, and the specific quality assurance mechanisms. Most evaluations are decentralized.

All independent evaluations are managed by ILO-EVAL certified evaluation managers and are conducted by a pool of internal evaluators or by independent consultants. Beyond capacity reinforcement, decentralized evaluations take place under a governance structure (hybrid system) that guarantees regular oversight by evaluation professionals, both in the regions and in HQ. Five regional evaluation officers provide direct support and oversight to decentralized evaluations and 14 Departmental Evaluation Focal points in coordination with EVAL provide real-time quality control that ensures decentralized evaluation ToRs, inception reports as well as the main evaluation document and related products are up to standard. The Evaluation Office at HQ maintains a close coordination of evaluation work with REO and DEFPs, and holds the responsibility of approving the evaluation report before submission to the donor.

The ILO's Evaluation Office conducts regular appraisals with external evaluation teams on the quality of decentralized (project) evaluations covering at least the minimum timeframe of 2 years. The latest appraisal found that there is no correlation between the budget for the evaluation and the quality, or the expertise of the team and the quality. However, there does exist a statistical correlation between time frame and the quality of the evaluation, the quality of the monitoring data, and the use of methodologies that are not specific to ILO's mandate.

Moving forward, ILO will focus on: developing an integrated planning mechanism to enhance the timing and sequencing of Evaluations; Working on clustering of evaluations; and the conduct of studies on specific methodologies relevant to ILO's specific mandate.

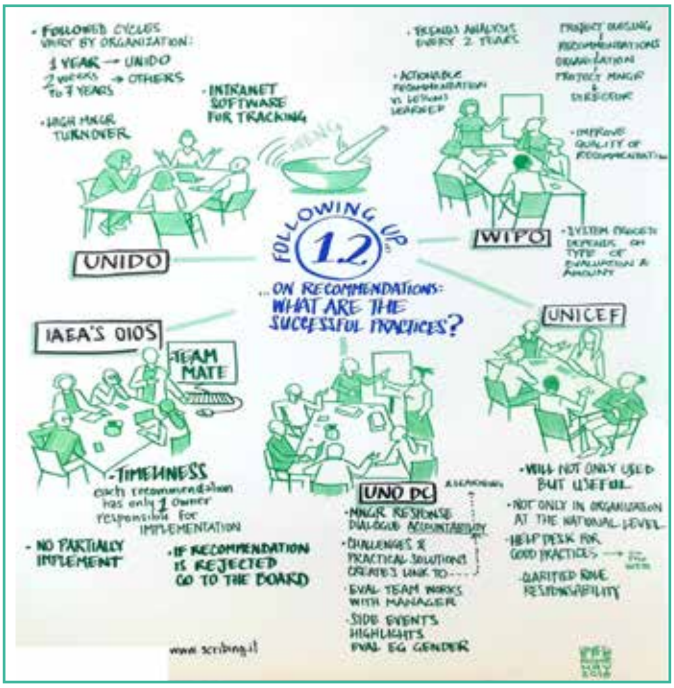
WFP case study on quality control during the evaluation design: framing and sequencing of evaluation questions

Diego Fernandez (WFP), shared a case study on WFP's assistance to a country struck by drought: the case study highlighted the importance of both framing and sequencing of targeted evaluation questions. "Good questions can help unveil the truth". A group discussion was organized on sampling, classification, and relationship with the government as it impinged on the neutrality of the findings.

Concluding remarks:

- Use of evaluations is not the same as usefulness
- UNICEF study on Influential Evaluations found that there was no correlation between high quality evaluations and influential evaluations.
- Timeliness on evaluations which foretell institutional shifts to alter the course of the institution. The potential of an evaluation to have that form of seminal impact is in selecting what is to be evaluated.
- At the entry level in projects, evaluation is required to see if a new CPF, quality of entry of new design projects, is needed.
- Clustering (SDGs and UNDAF) to respond to other elements, for results at a wider, broader level.

Session 1.2 – **Following up on recommendations: what are the successful practices?**
Lead convener: **Julia Engelhardt (WIPO)**
Collaborators: **Mark Keating (IFAD), Ada Ocampo (UNICEF), Michaela Berndl (UNIDO), Kamolmas Jaiyen (IAEA OIOS), Katharina Kayser (UNODC)**



This session focused on the ways managers/organisations grapple with recommendations resulting from evaluations, and on how evaluation offices try to ensure and track the implementation of recommendations. The session built on the sharing of practices and experiences, and on brainstorming on how to improve the utilisation of recommendations. Collaborators oriented the debate around two questions: i) what are the approaches used by various agencies? ii) what are the challenges and practical solutions?

Modality

Participants shared their expectations at the beginning of the session using mentimeter. The words with the highest frequency were: ownership, best practices and tools. There were 5 tables in a "world café" style. At each table, a different organization presented approach to monitoring and following up on recommendations, and facilitated a discussion around it.

Key messages

The key messages reflect the different models for follow up on recommendations, which were presented by different organizations, as well as the ensuing discussion around these.

UNIDO's following-up on recommendations approach

Method: Management Response Sheet and feedback sheet in annex on evaluation process itself
Tracking tool: Internal intranet
Timeline: follow-up one year after issuance of management response to recommendations

Overview of Recommendation Follow-up System

UNIDO follows-up on recommendations with a management response sheet. For each recommendation they assign one or more respondents that can be internal or external entities (for instance: a government). UNIDO also requests a timeline from the respondent or project manager to ensure that recommendation implementation remains time-bound. The management response sheet is distributed to all those responsible and the line managers – issuing the management response sheet to managers is a key lesson learned to create buy in and ownership.

One month after the distribution of the management response sheet, they request that the respondent report their level of acceptance of the recommendations (accepted, partially accepted, not accepted). Following this, a consolidated management response is issued.

If a recommendation is approved, the evaluation office publicizes this through their intranet, which is accessible by all staff. All responses recorded since 2006 are in this record. One year after issuance, they expect a second input to report on the action taken and status of recommendations taken (on-going, not implemented, implemented). This will give an implementation rate. They also include a feedback sheet in annex on evaluation process itself. UNIDO evaluation office uses this internally to answer critical questions surrounding whether the evaluation was timely and budget adequate.

The evaluation office reports to Member States on the status of implementation of recommendations. They also check the quality of recommendations as established through the manual.

Challenges identified:

There is a high Turnover of Evaluation Respondents: therefore, the **acceptance** rate of their respondents is very high (>90%) however the **implementation** rate after one year is a problem and much lower (35%). This is partly because there is high turnover in those responsible for implementing recommendations. Therefore the ownership of the recommendation is put into question. The solution that they identified is to address the recommendation to the office instead of the individual person. This ensures more continuity in the follow-up.

For partially or unaccepted recommendations officials have to include a summary explaining why the recommendation was not accepted, in order to ensure recommendations are being assigned to the right people. They also introduced "in consultation" and "in collaboration with" as options for reporting on recommendations to ensure that they were being sufficiently assigned to the correct people to make implementation realistic.

It is also important to create ownership in this process. UNIDO shared one case in which the respondents presented their recommendation implementation to Member States, which is an ideal scenario.

Following this overview the discussion covered the following issues:

- There was support around the table for this being a common challenge. UNIDO is now on the verge of extending their response cycle from 1 year to 2 years for implementation rate improvement. They have conducted extensive review of respondents to see what has happened with implementation, which has fed into this reform process.
- In FAO, they usually have a matrix for recommendations but the follow-up is erratic as they have no management response IT system in place.
- They found different cycles between organizations e.g. one was up to 7 years for the follow-up period to the shortest being 2 weeks with responsibility at the program office level which is out of the hands of the evaluation office.
- The group found that some organizations have no follow-up because the evaluation volume is too high and they are lacking capacity.
- Tracking systems were also discussed: the group agreed that it was interesting to see different software with different success rate in utilization – input may remain with evaluation office because inputs are not coming in from relevant offices. Some offices still use basic tools like excel to track recommendations.
- They also had a discussion on who are respondents and who should follow-up on recommendations to ensure follow-up.

UNICEF's following-up on recommendations approach

Method: UNICEF's model of Influential Evaluation – from actions to roadmap
Tracking tool: KPIs, quarterly dashboard, Help Desk
Timeline: A window of 1 or 2 years for the mgmt. response to be closed.

Overview of Recommendation Follow-up System

The group was offered a presentation on UNICEF's model Influential Evaluation. UNICEF started in 2009 with follow-up improvement discussions after an initial UNEG discussion on the importance of this, which later became mandatory in policy in 2013. This started as a best practice disseminated throughout later as having the guidance of the notion that it is important for an organization to have a follow-up function, which is critical to justify investment in evaluation.

Now they are moving towards making evaluations that not only are more used but that are also leading to changes: in other words, evaluations that are useful! They are not only talking about proposing changes within the organization (timeliness e.g. new policy on gender) but they are also looking to influence those beyond organization (e.g. agenda setting at national level). They are trying to demonstrate that evaluation does not just happen in the background and can be very useful for strategy development. For each step of the management response they put examples and good practices and clarify when the management response is needed and not needed.

When the evaluation is joint, it is advocated that recommendations should be joint in nature as well and specify as much as possible who is responsible across the joint bodies.

UNICEF tries to focus on clarifying the roles and responsibilities in the follow-up process. They are raising awareness for the head of office to genuinely take ownership and define what each role is at different levels of the organizations. Through several mechanisms, they are asking regional directors to report on the management response and on what changes this has led to. They now have a window of 1 or 2 years for the management response to be closed. They also have a help desk for good practices they share on the website (voices from the field and good practices)

In the template they do "agreed, partially agreed" as well as a column to implement recommendations (but there were many actions in here and some were not necessarily connected with recommendations so they found it hard to close and report progress towards implementing the recommendations). As such, instead of calling it "actions", they are calling it a "roadmap" - Guidance to ensure that there will be a logical sequence that will lead to the implementation of the recommendations. This will help to ensure that recommendations are closed as well as the possibility to show progress on their quarterly dashboard. This is working better – KPIs have improved on the implementation of management response.

UNICEF has an evaluation response tracking system to have management self-update and self-report. They are intending to do another review of management responses. Moreover, they do a lot of work reminding key management that the dashboard is coming to be updated and therefore to take advantage of their visits and contact with management to remind them about how they can use the findings of an evaluation.

Help Desk is considered very helpful. Inside is the name of the software they use.

Questions:

1. How do you address following-up on recommendations in the context of cross cutting thematic evaluations?

Depending on how complex they are, UNICEF usually communicates closely with the senior management who nominates a focal point for the overall management response team. This ensures responsibility not only at the top but also with line managers who will ultimately implement recommendations.

In the humanitarian office context where one has various reporting levels, it is the country office first which is responsible for following-up, followed by the country office.

2. How do you address unimplemented recommendations? Until when do you do follow-up?

If a recommendation is not implemented, a justification must be given and, usually, there is still pending issues.

3. What percentage do you have on your KPI scorecard for implemented recommendations?

They would say about 75% is a good baseline and they are achieving around that.

Following this overview the discussion covered the following issues:

- They shared lessons and changes instead of good practices in processes and tools to ensure that evaluations **will not only be used but USEFUL.**
- There was a discussion around a study that they are doing about how evaluations have led to different types of changes – not only in the organization but also at the national level.
- While it is good to have roles and responsibility spelled out, the quality of the recommendation will ultimately determine its success for implementation.
- Is there a way to ensure that managers will not only be put on the spot but also change something? They think there is no way to really fix this but they have improved the process by having added in the scorecard a management response area for compliance.
- There is the need to raise awareness that the management response is an internal process and tool that can fit into larger strategic plans and reports.
- They are promoting a more ‘road map’ approach with a logical sequence of actions that will lead to a meaningful management response.

IAEA’s following-up on recommendations approach

Method: Using Team Mate, each recommendation has only one owner directly responsible usually at the Director level, who will have to identify action to implement the recommendations.
Tracking tool: Team Mate (online auditing tool).
Timeline: Once program managers and others decide that a recommendation will be implemented, they go through the follow-up chart and follow-up twice a year. Because they report to the Director General and Member States, they have to issue a report every year on status of recommendations.

Overview of Recommendation Follow-up System

The evaluation function is under OIOD and consequently they follow the audit function in many ways. They use Team Mate, which comes from the audit function where each audit is treated like a project (how they track their time etc.). Once a report is finalized and agreed on recommendations, the evaluation office has the final say on the recommendations they want to use. After recommendations are agreed on and the final report is issued, they include with the report a summary of recommendations.

Each recommendation has only one owner (this is a restriction of Team Mate) usually at the Director level, who will have to identify action to implement the recommendations. Because they report to the Director General and Member States, they have to issue a report every year on the status of recommendations.

The process on recommendation follow-up is that although each evaluation officer is responsible from the start to finish, they have so many recommendations and units that they have decided that it would be more useful to have each evaluation officer follow-up on one recommendation so that they really know the background and context of each recommendation. However, sometimes the recommendations are not clear enough to agree on when it is actually implemented. If a recommendation is rejected right off the bat, it needs to have an accompanying rationale and report to the Board. This is therefore a big deal since it goes to the Board.

Once program managers and others decide that a recommendation will be implemented, they go through the follow-up chart and follow-up twice a year. After the responsible program managers have committed to implementing the recommendations, the team follows up on the evidence to prove it. However, since they have other work, there is a lag on checking whether evidence actually shows that a recommendation has been implemented.

Challenges:

A big challenge is timeliness because managers sometimes say the lag between reporting a recommendation as implemented and checking the evidence of this is too long. The evaluation officer has the power to say when they are satisfied with the evidence showing that a recommendation is implemented.

Sometimes recommendations are too broad (e.g. “should consider” status is not specific and actionable enough). Sometimes the recommendation is very action oriented (e.g. creates an action paper), but it is not clear whether the recommendation is actually addressing a problem. Therefore they close the recommendation which thus has a "not verified" status.

There is also attrition of staff where people move on, consequently a recommendation can become "closed" or "not implemented". This is a huge challenge because it ultimately means that it is rejected.

Most of the time, when a recommendation is given a "closed" or "not applicable" mention, it means that the money did not come or budget and program structure have changed.

Sometimes Evaluators follow-up on recommendations they do not agree with because they are not the authors.

There are discussion about providing another column in the summary-tracking sheet with a benchmark of where the evaluation office will actually be satisfied with the implementation.

They have over 100 recommendations per evaluation officer.

Following this overview the discussion covered the following issues:

1. How you should follow up will depend on the type of evaluation e.g. UNIDO has many project evaluations and ask for self-reporting but do not verify because this is not feasible due to the high volume.

2. What kind of quality control do you have of the relevance of recommendation follow-up?

This is linked to quality assurance and we need to have more of a discussion and consensus about this to have more agreement on recommendations prior to putting them into the formal reporting system.

3. Who validates the management response? WFP, for instance, have a "closed but not verified" option.

Now IAEA wants to have a grace period of 2 weeks where they come to talk to the Evaluation Office about the recommendations before they sign off. They are also in the process of revising the manual to formalize the process of providing a draft of recommendations and actions first that they both agree on and then collaboration continues after the recommendation is issued.

4. Do you have "partially agreed" recommendations?

No they do not partially implement recommendations because they take this from the audit practice where it does not exist.

5. Do you discuss the quality of recommendations together?

Not really but they want to start. UNIDO tried to do one on lessons learned last year but results were not great.

Conclusions and Take-Aways:

- Focus on roles and accountability
- Make sure that recommendations issued are implementable and actionable to ease follow-up.
- Timing: Be mindful of timing and staff changes to try to not let things drag on.
- Recommendations have a shelf life so be aware of this for relevance as things change with time.

UNODC's following-up on recommendations approach

Method: i) presentation to the project team and the management response at the same time, and ii) events and presentations to promote the results of evaluation recommendation implementation.

Overview of Recommendation Follow-up System

Once the report is finalized, they hold a presentation to stakeholders and present the report with the project team and management response. The project team responds right away and they have a dialogue around audience.

This provides a united front to show how evaluation and projects work together. It also created a higher commitment from the project team because it is done in front of an audience. However, although this works for shorter-term recommendations, it is not ideal for longer-term recommendations.

UNODC has side events to highlight evaluation (e.g. gender team and trafficking team). The project team will talk about how they are implementing evaluation recommendations, which is a positive experience for the project team and creates a link to accountability and learning. They discussed having more of a draft style debrief prior to the presentation but this would have to maintain independence. More are promoting a more qualitative way of reporting on implementation. They also discussed having more donor involvement.

It was observed that it was a challenge to confirm actual implementation and what was done.

They have an annual briefing where the management response and the evaluation are presented at the same time. A very participatory evaluation process is used (technical and strategic debriefing). M+E team in the field actually works with planning sessions of the project team and recommendations are included in the annual performance plan that is discussed during the yearly retreat.

Discussion:

1. Have there been any negative confrontations in the evaluation presentations?

Not really, it has been positive so far. Even if they reject the recommendations, it is still productive, but sometimes the discussion is around budget.

2. How do you balance the culture of learning and use vs. a more audit style top-down implementation?

In UNIDO, the majority of evaluations is made of project performance evaluations (which they do about 30/year). The only way they get learning across is to do more thematic and strategic synthesis reports and meta-evaluations to really hold management to account for this. The response they usually receive is that managers are hearing things they already know and cannot change.

To get the learning component, the buy-in from Directors at the top is needed. The challenge is to strike a balance between increasing learning while also being accountable.

Following this overview the discussion covered the following issues:

- Ownership and stakeholder engagement is critical to increase implementation and buy in.
- They gave examples of some of the ways they achieve this through debriefings and management response as well as side events they organise to provide positive examples of actual implementation of recommendations.

WIPO's following-up on recommendations approach

Method: Management response and survey following evaluation conclusion
Tracking tool: Team Mate
Timeline: Annually and biannually for Program and Budget reporting

Overview of Recommendation Follow-up System

WIPO provided a brief overview of the process to make recommendations more utilization-focused. The process was presented in six steps going from policy definition to reporting to governing bodies. There are several other steps in the process but a simplified version was shared with participants (see figure aside).

The discussions were then centred on specific questions from participants. The paragraphs below follow the flow of the discussion. Some of the questions included:

- How to make recommendations more utilization focused in the absence of such formal processes? Certainly, informal systems can also work but in the case of WIPO a system had to be put in place since the Evaluation Section has to report as part of its performance the status on the implementation of recommendations. Therefore, the Evaluation Section which is within the Oversight Division clarified formally the roles and responsibilities managers and the evaluators. Specially, since it is not the responsibility of evaluators to implement recommendations resulting from evaluation reports.
- They discussed more questions of whether this would work in other organizations.
- There were also concern as to what to do in cases when evaluators recommend to increased resources (HR or financial resources)? This sort of recommendations required departments such as HR or finance to approve this type of recommendations but this is easier said than done. What to do when this type of recommendations cannot be implemented?
- They use Team Mate, an auditing tool, and it would be interesting to do a trends analysis every 2 years to see improvement of recommendations implemented. However, WIPO indicated that there are also other tools on the market to manage the evaluation process.
- There are also other issues such as what to do when there are projects that are closing and have recommendation attached to them. In such cases, recommendations should go to the higher level, so from the project manager level to director level to ensure continuity.
- We need to be aware of the difference between recommendation and lessons learned. It is important to keep the number of recommendations realistic.
- Systems do help but what do you do with a huge volume of recommendation? How do you follow-up on 100 plus recommendations? Some offices like WIPO do only undertake strategic high level evaluations, there are other offices required to do project evaluations as well. This increases the number of recommendations an office would need to follow up. It would be important for these offices to explore realistic ways to keep track on the levels of implementation of recommendations but

offices with such a high volume of recommendations might need to undertake first a review of the quality of all their recommendations. They might need to assess what would be the best option for them without unnecessarily increasing the work load of existing staff.

- WIPO indicated that they have a team of 3 people in the evaluation office and it takes 5% of their time to follow on recommendations.

Session 1.3 – **As evaluation managers, we all face similar challenges: let's share and learn from them**

Lead convener: **Julie Thoulouzan (WFP)**

Collaborators: **Aurélie Larmoyer (FAO), Amélie Solal-Céligny (FAO), Ahmedou OuldAbdallahi (FAO), Sara Holst (FAO), Federica Zelada (WFP)**

This session allowed Evaluation Managers to reflect on common challenges in managing an evaluation and to brainstorm about several different strategies for addressing each of these challenges. These related, for instance, to steering the evaluation process in an effective way despite the fluidity and complexity of the environment, managing conflict within the evaluation team, or ensuring constructive engagement from key stakeholders.

Modality

Participants formed four different teams and took part to a best response game. For each question, they had a few minutes to brainstorm and determine a response to address a challenge that evaluation managers can encounter. A jury made of participants was responsible for selecting the most effective strategy.

Key messages

The paragraphs below report the best answers of the participants and their main debates.

Q1 - In the evaluation launch meeting, you find out that the report must be delivered earlier than what you had planned, because of corporate priorities in decision making. The evaluation team is concerned they will not have sufficient time to do the work. What is your best response to the situation?

Based on the assumption that the report submission date is not negotiable, participants isolated different solutions. First, the Evaluation Team could consider the possibility of **increasing team and/or resources**. Second, they should see if it is possible to **reduce the scope of the evaluation**. Third, it was agreed that the team should present preliminary findings and recommendations, making it clear that they are preliminary.

Q2 - The head of the programme being evaluated objects to one of the selected evaluation team members, arguing that she does not have the appropriate profile. What is your response?

It was decided that the best solution was to bring the case to the **evaluation reference group who would assess** the profile of the candidate and recommend the best candidate in a consultative way.

Besides, participants pointed out the necessity to be aware that the arguments given by the head of programme could potentially hide a bias against this evaluator (because of his age, gender, nationality or past experience). Hence, the importance of asking for a solid argumentation to support his/her assessment of the suitability of the team member. If the arguments are credible, alternate candidates should be considered. If the arguments are not credible, the person in charge should convey the final decision on the team selection explaining the rationale for his decision. He should make sure that he has strong references and records of past experience for this consultant.

Q3 - During the data collection interviews, the evaluation team does not have an appropriate behaviour. You either witness it or hear complaints from a government official on the ground. What is your best response?

A fitting behaviour would consist in calling for an urgent meeting with the Evaluation Team and Government to bring attention on the problem and find out what happened in order to mitigate. Besides, the presence of a senior staff to accompany the evaluation team in the future meetings with the government (to mitigate) can be considered. Finally, the person in charge should take the necessary measures and brief the team to avoid inappropriate situations to raise again.

Q4 - In the middle of data collection, one of the evaluation team members tells you she/he has to drop out. What is your best response to the evaluation team?

Participants agreed on the idea of consulting with the Evaluation Team member and try to understand the reason behind dropping out. Moreover, it seemed convenient to organise a consultation with the Evaluation Team to consider together different options such as **hiring** additional staff to replace the team leader, **sharing** the additional work created by his drop out, or **postpone** the evaluation deadline and identify a replacement.

Q5 - All the team members are fighting with each other. They don't want to talk to each other during the last week of the mission and inform you that they will not collaborate for the analysis and report writing. What is your best response?

The starting-point would be to discuss with team members individually to understand and try to negotiate and mediate. Secondly, the team leader could try to foster an interest based cooperation. In any case, participants agreed on the idea of preparing alternatives in case of worse scenario. For instance, the possibility for members to submit individual deliverables that the evaluation leader would consolidate was considered.

Q6 - During data collection, the evaluation team leader tells you that early findings suggest that the project shows systematic discrimination against some ethnic minorities. What is your best response to this situation?

Participants were divided on that question. The solution retained eventually was to probe further, understand, triangulate and seek evidence to make sure that the discrimination really happened as a first step. Then, the person in charge should discuss the problem with Management, check the organisation's policy in terms of reporting Human Rights violations, and, finally, include the findings into the draft report.

As per the ICSC Standard of Conduct, UN staff are expected to respect the dignity of all people without distinction; UN staff also have a duty to report through either the Investigation or Ethics office any breach to the agency's rules and regulations.

However, participants debated on whether to include the findings in the evaluation report or to share them with the ethics office and avoid covering these findings in the evaluation report. Some were partisans of a complete transparency, while others wanted to deal with the issue internally first.

Participants also shared some previous experiences. For instance, some faced a situation in which the report ended up not being published to which participants replicated that such a behaviour was risky since if the evaluation report is not published but information ends up being shared with the press, this can be a major reputational risk for the agency. In another experience, a participant explained that the case of discrimination in the evaluation report was recorded and reworded in a diplomatic language, while the ethics office was also informed (matter of credibility and transparency).

Q7 - The preliminary report has some critical findings, and one of the key programme partner sends a strong complaint email to your line manager, copying it to you and members of the evaluation reference group accusing you of incomplete and skewed data collection, and faulty data analysis. What is your best response to the situation?

It was considered that the best solution would be to organise a reference group meeting with all parties involved to discuss about data collection and analysis. Accordingly, if it happens that the complaints are correct, the person in charge should accept and modify the report. Conversely, if the complaints are not justified, then the report should not be modified.

Many participants commented on the necessity to consider the nature of the "partner" and to conduct, If needed, a stakeholder analysis focusing on which biases each stakeholder might bring.

Q8 - When the evaluation team approaches specific staff for interviews and document sharing, the staff have a defensive approach and don't want to share information with the evaluation team. What is your best response?

A first step would be to explain why it is important to share information with evaluators, what the purpose of the evaluation is, and ask the staff what are their concerns. Secondly, the staff should be explained that they have obligation to share the information. Eventually, the person charge should consider to talk to managers. Through the whole process, participants also noticed the importance of sharing communication that the evaluation team was coming ahead.

Concluding remarks

- All the scenarios discussed during the session were faced in real life by many of the participants in the room and were found to be highly relevant.
- The facilitator pointed out that there is no 'right answer' to any of these situations. Evaluation managers are working in fluid, challenging environments. In the same situation, different Evaluation Managers may take different approaches that might be equally effective.
- The following skills were highlighted: Evaluation managers need to be good communicators (active and empathic listening) and have effective negotiation conflict-resolution skills.
- A number of principles were recommended when working through challenges, based on these discussions: including impartiality but at the same time flexibility and ability to find solutions adjusting to an evolving context.

Session 1.4 – **How can we ensure a gender-focused evaluation?**

Lead convener: **Inga Sniukaite (UNWOMEN)**

Collaborators: **Felix Herzog (UNESCWA), Dawit Habtemariam (WFP), Catrina Perch (IFAD), Nicoletta Lumaldo (IFAD), Muge Dolun (UNIDO), Katharina Kayser (UNODC), Messay Tassew (UNWOMEN)**

In many ways the wheels of evaluation turn very slowly and as the profession has a long way to go to really integrate gender equality in the evaluation work. This session provided a forum to share and exchange ideas, good practices and practical experiences in integrating gender perspectives into evaluations, taking in consideration the varied capacities, focus area of work, challenges and the diversity of UNEG members. The session was also designed to challenge the participants to advance their thinking and work in this critically important area within the framework of the UN SWAP Evaluation Performance Indicator.

Modality

The session consisted in small-group discussions to foster the sharing of ideas, good practices, and practical experiences on how to integrate gender equality in evaluation work.

Participants were asked to identify **practical examples, possible solutions** and **innovative approaches** to the 5 challenges that were pre-identified as below.

Key messages

Challenge 1: Moving beyond sex-aggregation of data to evaluation of challenges in gender dynamics, roles and relations

To cope with this challenge, organisations should enhance capacity at design, implementation and monitoring levels – ask how much of the gender analysis has been based on collaborative/participatory approaches. At the strategic level – how does gender responsiveness get reflected and trickled down to what is implemented? How can the agenda be pushed forward? –, they should make good use of the theory of change.

Organisations should also focus on gender transformative thinking and harmonise the gender framework into the analysis.

It was also observed that when reporting back the analysis, organisations should try to move towards more qualitative means rather than focusing on the quantitative alone – having a balance between the two to more adequately represent situations.

Challenge 2: Gender responsive methodologies: How to evaluate power mechanisms in cost-effective ways?

The use of mixed-methods and participatory approaches was highlighted, to raise women's and marginalized voices with systematic and inclusive approaches. It is also important to communicate what evaluation is about with stakeholders: governments, UN bodies, academics, etc.

Participants also encouraged the use of UNEG guidelines and TOR's to understand and evaluate gender responsiveness. Moreover, in measuring the success of a programme in changing women realities and transforming gender relation, evaluators should ask if the root causes of gender inequalities have been addressed.

It is important to conduct thorough context analysis of gender in countries, and to include the use of feminist research. Attention must also be given to the inclusion of diverse stakeholders including sectoral ministries, gender machineries and women's organizations. Evaluators should use sex-aggregated data and adopt a broader and deeper analysis to get patterns that can be useful in gender framework. Moreover, social media data – Instagram, Facebook, Twitter – can be used.

They should also include GEWE as an evaluation criterion (whenever possible) and develop specific indicators with gender analysis.

Eventually, it was noted that evaluators should intend to apply feminist research in context analysis; and develop specific indicators with gender analysis.

Challenge 3: Highly technical area of work

Evaluators should intend to anchor all evaluations within the overarching goal or objective of the organization in relation to GE. They must make an active effort to engage more with women and include more women professionals/evaluators to increase participation and engagement with gender aspects.

There is also the need to identify gender issues in "technical" areas, i.e. think outside the box – many "tech" areas are human. However, there is also a need to refrain from tokenism – do not force it everywhere

On the other hand, organisations must have more female respondents or participants in projects by conducting more comprehensive data collection. Currently, there is gender imbalance in government, stakeholders and decision makers.

Participants also encouraged the adoption of a blend of quantitative and qualitative indicators within the evaluation matrix to assess gender equality outcomes, and signalled the importance of addressing the absence of gender analysis from a project design phase.

Challenge 4: Understanding the context to best evaluate gender

People must have a clear understanding that gender is not about women. It is about gender dynamics, roles and relations. The program context is not set to reflect gender – need to implement this at the design level. Organisations must consider each context that the programme is being conducted in.

In the case of the absence of change, evaluators should reflect on why change is not happening and try to understand the constitutional context, understand what the main gender equality considerations are and how well the intervention was designed to respond to these; they should also undertake stakeholders analysis and design methods to reflect and/or address stakeholder diversity and needs

Challenge 5: Lack of integration and analysis of GEWE into programme design and its implication on gender-focused evaluation

Organisations should foster participatory learning throughout the process to minimise resistance, work with evaluation teams that have background in gender responsive evaluations that can respond to the issues, or use priorities and guidance of organisations, i.e. Gender marker, organisational strategies and policies.

Moreover, they should capacitate/better support decentralised stuff – tools, guidance and trainings so they can better integrate the gender dimension. The importance of gender policy at corporate level was also emphasised. On that topic, the networking can be a good instrument as there are innovative processes and tools used by different organisations that others could learn from.

Concluding remarks:

- Creating demand within organisation for gender responsive evaluations is a shared responsibility of the evaluator and programme personnel.
- Improve the quality of policies, normative work and programmatic areas to better include the gender aspect.
- Engage with the UNEG Human Rights and Gender Equality Working Group.
- Apply and use UNEG guidance, tools and other documents on gender responsive evaluations.
- Engagement in the Peer Learning Exchange facilitated through the UNEG HR and GE Working Group.
- Use the UN System Wide Action Plan on Gender Equality and Women's Empowerment Evaluation Performance Indicator (UN-SWAP EPI) as a key instrument to push the integration of gender perspectives throughout the evaluation processes and products

Session 1.5 – **The role of donors in evaluations: how should we involve them and how to manage expectations**
Lead convener: **Amir Piric (UNESCO)**
Collaborators: **Thuy LE (UNIDO), Ahmedou OuldAbdallahi (FAO), Luisa Belli (FAO)**

This session was an opportunity for the Evaluation Managers/Officers to exchange on their experiences in dealing with donors in general during the evaluation process. Participants reflected on the common challenges they faced managing their relationships with donors and discussed upon good practices to address them. The discussion also encompassed ways to engage donors so they become strong supporters and advocates for evaluation. Eventually, the session reflected on ways to leverage on this donor support to further promote a results-based and evaluation culture within our organisations. Donors' involvement both at macro (organisation) and micro (evaluation) level were discussed.

Modality

UNESCO shared an informational note on the organisation's experience with MOPAN and key lessons learnt, following which the conveyer opened the floor for other organisations to share their experiences with MOPAN. The second part of the session entailed presentations by FAO and UNIDO on a broad range of experiences with donors during the course of evaluations.

Key messages

Description of the relation between UNESCO and MOPAN

UNESCO welcomes the general principles guiding the approach of MOPAN, which represents a coordinated assessment of 18 partners, thereby fostering system-wide coherence in line with Paris Aid Effectiveness principles, and (potentially) reducing transaction costs for assessed organisations compared to bilateral assessments. The institutions which have engaged with MOPAN include notably UNICEF, WHO, UNHABITAT, WFP, UNWOMEN.

Strengths in the methodology include the variety in sources applied against the MOPAN indicator framework and the involvement of Member States of the assessed organizations in the overall assessment process through the designation of an institutional lead which acts as 'go-between' the MOPAN Secretariat and Member States. Assessment has been conducted in a professional, collaborative and transparent manner, with good working relationships between the MOPAN Secretariat and the UNESCO team.

Suggestions for improvement of the MOPAN methodology

The importance of consultations in the preparatory phase and the definition of the overall timeline of the assessment and its various phases was stressed (e.g UNESCO has had a change in leadership which will certainly lead to reform in many areas, hence the results of the assessment may become 'outdated' in short period of time).

Another aspect concerns the fact that it is not clear how ongoing high-level reform efforts and their implications - the SG reform proposals for a repositioning of the UN Development system - can be reflected in the assessment.

It was signalled that the selection of pilot countries for each MOPAN cycle should be done in coordination with each Organization - this would increase the relevance of findings for organizations that have a diversified field network.

The reflection of normative and standard-setting work in the indicator framework could be improved. Moreover, KPIs 9-12 do not seem to apply in equal measure to all organizations - the possibility of adding organization-specific segments here should be considered.

Technical difficulties with the stakeholder survey can be observed: Field Offices/country contacts reported challenges in terms of internet accessibility and language constraints which puts in question the representatives/validity of the exercise.

Eventually, MOPAN Secretariat could be more explicit about the themes/questions that will guide the interview week to ensure informed selection and preparation of interviewees and to make clearer the strategic intent of the segment and how it relates to the overall assessment.

Lessons from UNESCO Evaluation Office with regard to KPI 8: Evidence-based planning and programming applied

- Evaluation Policy used as a key reference in providing evidence to all 7 Micro-indicators;
- Sufficient evidence coming from corporate evaluation products and processes;
- The quality of evidence coming from the decentralised evaluation is uneven;
- The extent of the background documents and interviews conducted etc. is vast and quite comprehensive;
- MOPAN external reviewers were receptive to comments and willing to accept critiques.

UNESCO's experience with MOPAN

The benefit of MOPAN is that peer-review can be complemented with MOPAN. Key performance Indicator 8 is used as a key reference document. MOPAN selects 8/9 member states for an in-depth analysis.

MOPAN current indicator framework may not be best suited to demonstrating the positive results claimable by an organization like UNESCO which undertakes substantial normative work. But MOPAN's current methodology is still evolving: a number of issues are surfacing with the Stakeholder Survey. One of the observations which was made during the deliberations in the session is that MOPAN's current assessment doesn't have an interview protocol.

Other organisations' appreciation of MOPAN

1. UNWOMEN elaborated on the potential for a linkage between a MOPAN assessment and the relationship with resource mobilisation
2. WHO recounted their experience where a peer review of the evaluation function last year fed into KPI 8 of MOPAN, and stressed that the exercise undertaken by MOPAN is a not a duplication. However, a MOPAN assessment exercise takes as long as 18 months, hence, timing is an issue.
3. WFP: completed by 2013, managed by the office of evaluation; the current MOPAN is managed by the division on performance management; it makes strong use of background reports.
4. UN HABITAT shared its experience with MOPAN by indicating that their assessment exercise focussed on 2 program areas: management and performance; they use a Theory of Change, inception report, and a stream of evidence. Has wide buy-in.
5. Should the UN be involved in the development of the MOPAN methodology? This is related to the lack of integration of the SDG agenda etc.

Specific Evaluations and Engagement of Donors

FAO's experience

FAO shared an experience evaluating the African Solidarity Trust Fund which is financed by two African countries, namely Angola and Equatorial Guinea. Since the 2 donors have not been involved previously in evaluations conducted by FAO's Office of Evaluation, it was important to engage them from day 1 to explain the evaluation process and understand their expectations from the assessment. He indicated that the evaluation commenced by clarifying expectations with the Steering committee which included representatives of both countries. Throughout the process, the evaluation team updated the Steering Committee on the progress and received comments on the key evaluation products, including the concept note, ToRs and draft evaluation report. The experience that he shared made it evident that, just like in the case of other key stakeholders, engaging the donors during the evaluation process is useful in managing their expectations and improving communication. It is also an effective way to sensitize them about the evaluation culture in the hope that they will be strong advocate for evaluation.

UNIDO's experience

UNIDO shared an experience where a thematic evaluation of UNIDO's Partnership with Donors was conducted. Although the exercise initially met with some resistance, the conduct of the evaluation in a transparent manner eased all concerns: donors were involved in key steps of the evaluation, selection of donors was undertaken in a systematic manner, key evaluation products were shared with key stakeholders and donors for comments, interviews and surveys of donors undertaken both with representatives from the Permanent Missions and from the capital and etc. Very positive findings surfaced from interactions with donors who highly appreciated the opportunities to provide "customer satisfaction feedback". Some interesting insights emerging from this exercise were: dialogue with donors was essential; evidence and independence are key to enhance evaluation credibility with donors; and donors should be handled in a case-by-case basis depending on their interest and operational preference.

FAO's experience 2

Donors are themselves a part of the evaluation, i.e the *evaluand*, and they influence the course of the evolution of the work of the organisation, but they are also instrumental in altering the course of the implementation of the project through late release of funds, because some funds are never released as promised etc.

It was remarked that donors are not homogenous entities, and have varying degrees of interest in the process, regarding the conduct and the results of evaluations. For instance, some colleagues noted that in their experience donors from Gulf States have shown limited interest in evaluations.

UN Women raised an interesting question on donor-mandated evaluations: how do you contain donor-mandated evaluations (mostly project evaluations), and how do we reorient donor-driven evaluations to include more strategic/thematic/country/regional evaluations? One strategy would be to include cluster evaluations which fall under a similar theme. Moreover, DONORs have the OECD-DAC platform to get donors to think about what is relevant to them, similar to how evaluation agencies have the UNEG. FAO shared the institutional experience of the organisation in instituting the trust fund for evaluations which eliminates the need to repeatedly deal with donors to commence evaluations. Some other organisations echoed the same sentiment.

Session 1.6 – **Procuring UN evaluators externally: learning from current practices and suggestions for improvement**

Lead convener: **Andrew Fyfe (UNCDF)**

Collaborators: **Luisa Belli (FAO), Isabel Suarez Garcia (UNWOMEN), Robert Stryk (UNRWA)**

This session addressed the difficulties that UN evaluation systems meet when it comes to hiring and working with external consultants. It is observed that procuring credible, expert evaluators is not easy. Moreover, there is evidence of growing frustration from some suppliers of evaluation services of the difficulties in working for international commissioners of evaluation, ranging from a lack of clarity around what is expected from evaluators to a lack of time or budget. The session has two objectives: i) to start a conversation between UNEG members on good practices in designing and managing the procurement of our evaluations within the UN system; ii) to consider possible responses to the broader evaluation community on what we can all do to improve the practice of seeking or procuring our evaluation experts.

Modality

The sessions was divided into four parts, each corresponding to a different stage in the evaluation management cycle, though with a focus always on the procurement or contractual arrangements with the evaluators in line with the broader theme of the session. Participants self-divided into the four groups before returning to Plenary to present their main conclusions.

Overarching question: Are UN systems for procuring external evaluators working well?

Key messages

Background

While many UN evaluation entities depend wholly, or in large part, on external evaluators to ensure the independence, credibility and usefulness of their evaluations, it is not always straightforward to procure these evaluators, and this for a variety of reasons:

- Difficulty in ensuring a reliable supply of credible evaluators who provide expertise in both the technical areas being evaluated as well as in up-to-date evaluation methodologies as well as other evaluation competencies such as skills in GEEW, and regional and local knowledge of areas being assessed.
- Supplier base of reliable/credible/expert consultants and firms appears to be shrinking in some countries. These firms are sometimes unavailable/ have multiple engagements with other evaluation/are unable to complete evaluations because of multiple commitments to others.
- Budgets for evaluation are sometimes decreasing while the costs needed to conduct good quality evaluations in line with UN standards are increasing

It of course takes 'two to tango' in any contracting arrangement, and from their side independent evaluators sometimes express frustration with the commissioners of international evaluations (including UN agencies) for the following reasons:

- Evaluation budgets can be smaller than what is required to meet the requested scope and quality standards of the evaluation;
- Procurement processes can take longer than expected, particularly when compared to the short timeframes that evaluation advertisements are often published for;
- Once contracted, too little time can be given for the evaluators to mobilise their teams, conduct the evaluations and oversee participatory write up and finalisation stages in the process;
- Evaluation TORs can be overly prescriptive with insufficient space for evaluators to propose their own approaches to the evaluation; at the same time the more that we demand from bidders, the more expensive it can be for them to comply.

1. Choosing the initial procurement modality

There are lots of different modalities that UN evaluation entities can use to procure their evaluators. Each can have its merits and potential drawbacks depending on the type of evaluation being procured and the organisational environment in which UN evaluation entities find themselves.

Some of these modalities include: initial expressions of interest, leading to long-term agreements with short-listed evaluation firms or university or research entities; request for proposals with individual evaluation firms; roster arrangements for speedier procurement of individual evaluation consultants.

Advantages/drawbacks of different evaluation modalities

Type of procurement modality	Advantages	Drawbacks
Expression of Interest, followed by additional step	Clear signal to evaluation market of intent to proceed; allows more precise targeting of specific evaluation profile sought	Not very time efficient as a two-step process. However, can lead to a very effective evaluation process if the 'right' firm is identified.
Long-term agreements with specific firms or research institutes	Can speed evaluation procurement up particularly when multi-annual evaluation requirements are known. Helps evaluation bidders know level of effort ahead of time which helps with planning.	Can restrict opportunity to go to market, particularly for interventions or evaluation designs that are innovative or very specific in their requirements
Request for proposals	Workhorse procurement modality for many, can ensure quite precise specification of what is sought	Depending on requirements of bidders, can be onerous to adhere to on behalf of evaluation bidders
Evaluation rosters	Can be an efficient way to procure individual consultants quickly	Not always a very precise procurement tool, and often requires additional procurement steps before proceeding.
Contracting of individual evaluators	Another workhorse procurement modality for many – can be very specific depending on evaluation designer	Can be quite time consuming to procure. Can also require additional management throughout the process in cases where consultants haven't worked with each other before or where evaluators don't know the UN system.
Contracting of additional evaluators to join established evaluation teams from Evaluation Offices	This can add expertise and independence to existing teams.	Can add more time to the process.

2. Designing the chosen procurement modality to incentivise evaluators to respond

As experience varies across the different UN entities, participants across different UN entities isolated the following elements as good practices.

Depending on the maturity of the evaluation market that is being explored, it is important to clearly define the objectives of the evaluation, and the evaluation methods and tools that should be deployed in order to meet evolving quality standards for evaluation reports in the UN system. In doing so, we must consider the various merits of more detailed, more prescriptive evaluation designs in cases where this serves as a quality assurance tool versus more open-ended, less prescriptive designs where evaluation capacity within the marketplace is more assured. Taking care at the same time not to make the bidding requirements overly onerous for evaluators to reduce the cost in bidding.

It is important to consider carefully the amount of time that is needed to conduct the evaluation and be transparent about the estimated level of effort required across clearly defined stages of the evaluation (e.g. inception and finalisation of the evaluation approach, collection and analysis of secondary and primary data - including visits to the sites of the interventions and interviews with key stakeholders if appropriate - while taking care not to infringe on the requirements for fair and open competition e.g. by providing an estimated number of days, rather than a clear budget for the evaluation, rather than a clear budget).

It is also essential to provide transparent guidance on how evaluation bids will be assessed, with clear weightings between different assessment criteria: e.g. experience of firm, expertise and experience of evaluation team against clearly-defined expert profiles, quality of the methodological proposal, quality of answers to interview questions for shortlisted firms etc.

3. Selecting the firms

This relates to processes that evaluation entities use to make the procurement decision. Elements of good practice can include the following.

Participants stressed the importance of ensuring that selection panels are constituted carefully in line with the evaluation and technical expertise sought for the specific evaluation being procured (UNDP/UNCDF rules require at least three panel members with different skill sets). They also support the principle that the format of the selection process (e.g. initial technical assessment based on desk review followed by interview of shortlisted evaluators, including detailed scoring matrix) is clearly stated in the initial document advertising the evaluation.

Moreover, when working with a new evaluation firm, it is important to request evidence of previous evaluation work as well as detailed reference checks (procurement divisions can provide useful guidance to support this).

Eventually, for unsuccessful bidders, it can be useful to provide a telephone debrief, pointing out areas for improvement with a view to encouraging bidding again for new evaluation opportunities. At the same time it is important not to give away any information about competing firms to respect UN competition and procurement rules.

4. Managing the evaluation from beginning to end

This refers to the contractual arrangements for managing the evaluators from the start of the process to the end which can also determine the willingness of external evaluators to work again as UN evaluators.

It seemed important for participants to be as transparent as possible about expected deliverables and the timeline for them to be delivered from the very beginning. This helps evaluators to plan their work across the duration of the year. The importance of planning regular check-ins and milestones for deliverables to set clear performance expectations was highlighted. The fact that evaluators understand what is expected of them professionally and ethically must be ensured. Organisations should also provide relevant guidance documents (inception reports and evaluation report templates) as well as the expected standards that the evaluations will be assessed against (e.g. the standards for UNDP decentralised evaluation; also the standards for an evaluation that meets UNEG gender equality and women's empowerment requirements).

Finally, participants noted the importance of managing carefully the relationship between evaluator and evaluand throughout the evaluation, making full use of the evaluation unit or commissioner as neutral player and custodian of the broader evaluation standards (impartiality, credibility, utility) that need to be respected.

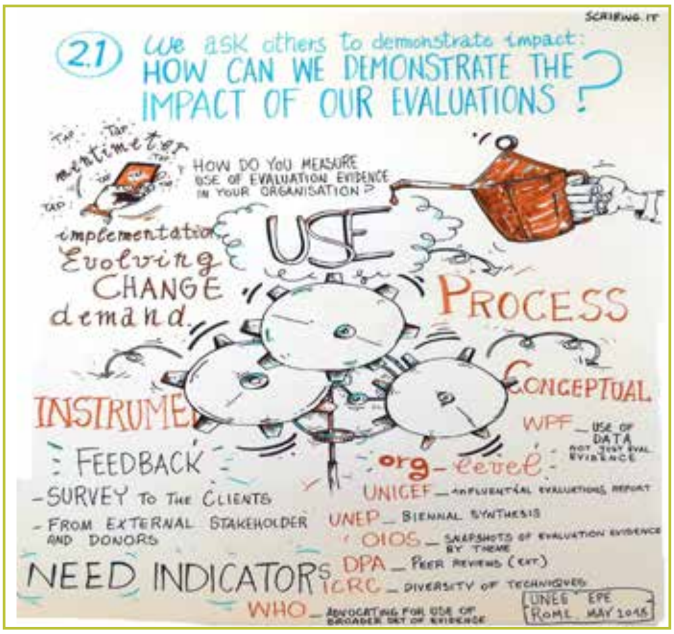
Concluding remarks

Participants agreed that there could be room for further work, perhaps within the professionalization workstream of UNEG, to set out and document good practice in this area given its importance in getting evaluations off to a good start. We agreed to follow up with the UNEG management to see if there would be space and interest to do this.

Session 2.1 – **We ask others to demonstrate impact: how can we demonstrate the impact of our evaluations?**

Lead convener: **Deborah McWhinney (WFP)**

Collaborators: **Amélie Solal-Céligny (FAO), Marta Bruno (FAO)**



This session proposed to define the 'impact of evaluation' as the 'use of evaluation'. The latter endorses three main uses among which two were discussed: i) conceptual use, which refers to the evolving conceptualisation and understanding of those who design strategies and programmes; and ii) process use, which refers to the changes to implementation or programming resulting from interactions between the evaluation team and key stakeholders. The following questions led the discussion: i) what entry points can you identify to increase the use of evaluation evidence in your organizations? ii) what are the challenges that you have faced? iii) how have you used the evaluation process to increase the use of evaluation evidence?

Modality

Attendees gathered in the Ethiopia Room and the conceptual framework was discussed – namely, defining 'impact' as synonymous with 'use' of evaluation evidence. 3 types of 'use' were defined: instrumental, conceptual and process use. It was agreed that there would be a focus on conceptual (influencing strategies and programmes) and process (changes to programme results due to interactions between the evaluation team and stakeholders).

Participants were then prompted to connect to the Menti website to respond to the following question: *How do you measures use of evaluation evidence in your organization?* Responses were shown up in the room's screen. Responses were pretty diverse: an overwhelming number referred to the follow-up to evaluation recommendations; a few respondents indicated they were not sure of the way evaluation evidence was measured in their organization while other responses related to the following tools used to measure evaluation evidence: tracking the number of evaluation downloads, requesting follow-up reports from evaluands and discussing in meetings of government bodies.

The room was then divided into two working groups working for about 20 minutes:

- One group on conceptual use
- One group on process use.

At the end of the session, summaries of each group's discussions were shared in plenary session. No follow-up actions were proposed.

Key messages

Group 1 worked on conceptual use – specifically, the initiatives and processes in place to measure and report on use in their organisations. The following inputs/ suggestions were provided by various agencies including: integrating a review of evaluation evidence systematically in programme review processes; evaluation budgets embedded at regional and country levels; tracking 'use' formally (KPIs/ corporate reporting) and informally (e.g. "Influential Evaluations"); external reviews of the evaluation function; and, responsive and engaged management response system (more than mechanistic).

Organisation	Initiative
WFP	• Systematic review of all new Country Strategic Plan documents (all 83 WFP Country Offices) with comments and suggestions on the use of evaluation evidence to enhance programme design, as well as evaluation planning and budgeting • OEV participation in strategic Programme Review Process meeting chaired by the AED • Working to understand how programme colleagues search for evidence/data to inform their work; influencing behaviour change
UNFPA	• Quality assurance of Country Portfolio Evaluations • Participation of OED in Programme Review Committee
UNICEF	• Costed evaluation plans with each Country Programme • Country Portfolio Evaluations mandatory in each CP • "My Impact Logbook" – note-taking of informal, organic feedback on the use of evaluation evidence • *Influential Evaluations* report • Need to go beyond evaluation evidence and data to consider the holistic context in which country-based programming is done
UN Women/ILO	• Budget for evaluation included as the country and regional level • Regional Evaluation Strategies (UN Women)
UNFPA	• Quality assurance of Country Portfolio Evaluations • Participation of OED in Programme Review Committee

UNIDO	• Management formally briefs Member States on their response to evaluation recommendations • Tracking statements by senior managers and analysis done
OIOS	• Assessment of the evaluation function in UN funds/programmes every 2 years – why have things improved or not? • Snapshots of evaluation evidence by theme
UN Women	• Question on use in Annual performance Report • Meta-analysis of evaluation evidence conducted
UNEP	• Biennial synthesis – notes changes in ratings for evaluations over time
DPA	• Peer review process
ICRC	• Process and product important for learning and use • Diversity of techniques needed to enhance engagement • Engagement by Divisions of Evaluation in the evaluation process to ensure learning
WHO	• Weak evidence base has led to advocacy for use of a broader set of evidence/data

Group 2 worked on process use – specifically, the **evaluation process** and how use can be enhanced between evaluation teams and stakeholders in the evaluation process. The group started by asking **how to do to get feedback from clients?** In some cases, they do a survey at the end of each evaluation, and summarise the information once a year (within annual reports for instance). The group reported that typically there is lot of frustration from the users (the evaluands) when they feel they cannot share feedbacks.

External reviews of the evaluation functions, performed each two years, are also an opportunity to have a third party communicating with the Evaluation offices and the clients.

In one case, an Organisation reported it has received requests from program units on specific technical topics; in this situation, a team collects all findings from evaluations on a specific topic and then share them with the technical unit. The group indicated **that Evaluation offices should create this type of demand with technical/operational teams** so that this interaction could happen more.

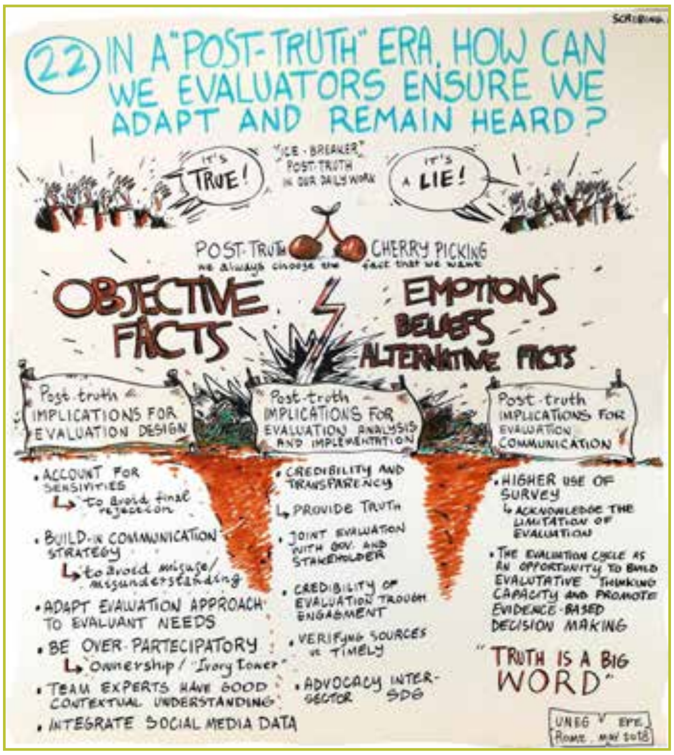
Another initiative is monitoring the number of project managers who confirm that they have applied the project recommendations. After each big evaluation, a UN entity go back to the Evaluand to check the key changes that were implemented after the end of the evaluation.

Finally, the group reported it would also **need some indicators on how the quality of programming has improved as a result of the evaluation but that it does not have guidance on this.**

Session 2.2 – In a "post-truth" era, how can we evaluators ensure we adapt and remain heard?

Lead convener: Adan Ruiz Villalba (WIPO)

Collaborators: Aurélie Larmoyer (FAO), Patricia Vidal Hurtado (ILO)



This session addressed the difficulties evaluators encounter regarding their credibility and their influencing power in a post-truth context. Such a context implies that evaluators may be seen just as uncertain as other information professionals, or that they may be challenged by more attractive narratives. Participants discussed and reflected on how the evaluation profession in the United Nations can provide an answer to alternative facts/post-truth in the public discourse, and on how to support decision-making in the international arena during the era of the SDGs.

Modality

The session started with an icebreaker on lying and truth. A participant was invited to say one statement that is he pretended to be true and participants had to determine whether it was a lie or a truth.

Then the topic on post-truth was introduced and connected this to SDGs. Participants were invited to take part in a role play situation to do demonstration about how post-truth affects daily work.

Participants were then split up in 3 groups and try to find solutions and exchange experiences related to situations where they had to fight basic lies and alternative narratives that are not evidence based. Discussions concerned different phases of evaluation, namely design, implementation and dissemination of findings. The session ended with a comeback into plenary to share final thoughts.

Key messages

Role-play about how post-truth affects the work of the participants

Role-play situation: a project in the UN that has a huge budget and investment has been rolled out with thousands of people and only now are the evaluators presenting findings to their Director. The team played out a situation where two evaluators presented the findings and evidence to the director who "took note" of the findings but did not sufficiently take it into account as he had "alternative evidence" he wanted to rely on instead for personal or political reasons.

Presentation of Theme Definition

Post-truth is relating to emptions and personal beliefs to defend what one believes to push away facts. It is something that is used to dismiss establishment politicians seen recently in politics of some countries. There is the worrying pattern that expertise is dismissed compared to what is heard on social media. It seems that now anyone can become someone with an important opinion and opinions have become facts.

This is relevant to our profession of evaluation because we are the experts who are now being dismissed when facts are being put in question. As such, what does post-truth imply for us and how can we respond as evaluators?

As will come out in the discussion today, we should react by holding on to standards and good practices again and by reiterating why we have good practices. We will also be contextualizing this in the 2030 context of development agenda to examine how critical thinking and evaluative evidence can be used to promote fact-based decision-making.

Group Reflections:

1. Design– FAO

They came from the angle that the results of evaluation reports are to be challenged by alternative facts so evaluators need to step up to the plate in terms of becoming better influencers. The team came up with several important points for how evaluators can be better influencers in the design stage, reported below:

Account for sensitivities by knowing the context to avoid final rejection: it consists in making sure that the context was well-analysed; and anticipating the possible end stage discussions and the rejection of findings, which could be biased perhaps due to sensitivities that are personal or political. This may build wall between the evaluator and the evaluand and make them inclined to ignore or look at "other evidence".

Build in a communication strategy: from the design stage it is critical to have a budget in a communications strategy (e.g. the briefs for decision makers that do not have time to read a long report and extra sessions and events to promote and market the evaluation so it can be deemed as useful). Also it is important to have a communication strategy to avoid misuse of the evaluation by possibly be taken out of context.

Use evaluative thinking – use the design stage to select good questions to guide evaluations: it consist in not just blindly use DAC criteria and Norms and Standards but listen to evaluand and shape the evaluation questions to context using evaluative thinking.

Be over participatory – build in ownership and get out of your ivory tower: set grounds for collective critical thinking and participation.

Use good qualitative inquiry through storytelling: team experts have the ability to leverage qualitative stories to further the contextual understanding and persuasiveness. Qualitative inquiry is good because it appeals to emotion. Stories are power.

Integrate social media data; while acknowledging the noise.

2. Implementation – ILO

Q1: Engaging stakeholders – consultative fora (learning process). The process of an evaluation is often times more important than the result of the evaluation itself. What can evaluators do to increase the credibility?

A solution is to go back to the basics and make evaluation participatory, e.g. what are the conditions that should be in place to facilitate the use of evaluation findings?

- Evaluation Reference Groups: national and agency level reference groups to help provide context and information for evaluation.
- Preliminary outcomes validation - Through a briefing session of assumptions and preliminary results. They assume results achieved and test their hypothesis. This is a bit like a theory of change that is linked to an evaluation matrix.
- Use national data for evaluation – Show how evaluation can link to this data to build credibility and use existing national capacity.
- How to promote transparency, credibility, and independence?

The real question is whether all these things work (going back to the basics and relying on norms and standards etc.)? Is there evidence that this works? There is no silver bullet but it does help to come back to the basics, e.g. put differences in data as a limitation to the evaluation report so as to maintain credibility.

Q2: Verifying source reliability. Is there a possibility to increase the reliability of evaluations without impacting timely delivery of evaluation products?

Good practices were identified to improve reliability of evaluative information: the research Analysis Unit that tries to provide readily available checked data sets that are palatable in order to broaden the evidence base (WFP, UNFPA); it is important to organize ourselves well: storing key basic data and documents to share with external consultants and drive evaluations as a base to always go back to as the core source. This is more an administrative suggestion to show how some organizations store and organize data.

E.g. FAO used to collect earth data to gain an idea remotely what kind of results they should expect and this is a rare case and very context based.

We can make members of the evaluation team aware of the possible bias in internal and external data to combat credibility issues. Indeed, it is critical to be clear on limitations that the evaluation team experienced in data gathering.

Q3: Evaluation report and recommendations (sound product and responsive to diversity of constituency needs): What can evaluators do so evaluative evidence (and not only statistical information) is used to support conclusions on development progress and SDGs?

In a period where we are shifting from donor driven programming to nationally driven programming, we have to ensure that we are building national capacities through decentralized evaluation. However, this causes problems from the quality assurance point of view.

Participants stressed the utility of joint evaluations in that context – to share responsibility and promote credibility and independence

It is also critical to clarify roles and responsibilities on who is meant to evaluate progress in achieving the SDGs at national level

Participants also signalled the utility of inter-sector advocacy – taking an inter-sectorial approach. UNDP has experience bringing together government and national evaluators to share experiences and practices.

3. Dissemination and communication – WIPO

Truth is a big word for evaluation. Evaluators try to provide different perspectives to the extent possible.

They have to be mindful that often the most important moment for senior management is when the draft report is given to them and when they acknowledge that this document will be public at some point. This is a key moment to influence through evaluation.

The reliability of data and sources of data are critical. E.g. the group gave an example of the use of HR data, when the program manager waited until last minute at the draft stage to correct the data that the evaluator had. Managers often wait to the end to say that data in the analysis is wrong. How to avoid this?

It was noticed that evaluators can increase credibility by linking to stakeholders for buy in. The group also entered a discussion on how to address track-changes. Some use track changes and some use table of comments → how to address comments and check whether comments have been addressed is very important!

UNICEF brought a discussion on trade-offs between real time and quality that depends on the level of involvement of stakeholders → There is a need to be aware that one needs quality to be able to influence and there is the danger that if we are too quick in conducting the evaluation to get it out fast, it may not be favourably received as quality may be compromised.

E.g. 30% of World Bank documents have never been downloaded from website

Call to action: There was a call to action by the EPE session participants to work together as an evaluation community to deal with this phenomenon of post-truth and alternative facts together. Eval Partners is an example of such an initiative – a flagship program tool advocacy campaign to counter balance post-truth era messages (UNICEF).

Session 2.3 – **Cost of evaluations: how to rationalize the use of resources for evaluations?**
Lead convener: **Amélie Solal-Céligny (FAO)**
Collaborators: **Ahmedou OuldAbdallah (FAO), Aurélie Larmoyer (FAO), Sara Holst (FAO)**

Evaluation Managers often have to justify the cost of evaluations to Project Managers who do not see the rationale for a high budget. Evaluators are sometimes challenged by evaluation users on the use of resources, related either to the time spent in the field, or the size of the evaluation team. On the other hand, evaluators might face requests from donors who require specific impact assessments that may increase the budget of the evaluation. Besides, evaluators are sometimes confronted with people who argue that there is no need to spend resources for an evaluation office when there already is an audit function in an organization. During this session, participants shared their experience related to justifying or adjusting the cost of evaluations or the evaluation function in their organization.

Modality

The session consisted in a role playing game staging four different characters debating about the costs of evaluation. Participants are invited to intervene directly or through mentimeter. The first debate opposed an evaluation manager (EM) to a project manager (PM). The second debate opposed a director of evaluation (DE) to a director of cabinet (DC).

Key messages

1st debate: staging an evaluation manager (EM) and a project manager (PM) discussing about three different topics

1. Budget of data collection:

From the PM's point of view, the data collection seems expansive and redundant as it was already done during the project. According to the EM however, evidence from beneficiaries and talking to people are critical.

After debating, it was commonly recognised that the EM should use available data as much as possible, to free funds that could be allocated to other missions. However, those data must be validated and new data must be produced if necessary. The final decision is up to the EM.

2. International vs national consultants

The PM judges international consultants too expansive. For the EM, national consultants should be prioritise to the extent that they have the necessary technical skills and independence. Bad previous experience with national consultants can be an argument against hiring them.

Through this debate, it has also been remarked that inter-agency joint evaluations should be encouraged.

3. Stakeholders meetings

The PM does not see the point of organising expansive workshops. He argues in favour of an online workshop. He also wants to participate in choosing who will be part of the workshops.

For the EM, workshops are highly important. It is determinant that people meet and share their feelings about the project. EM must have freedom in choosing the participants to create a proper discussion.

Regarding donors, it has also been emphasised that workshops are an opportunity for showcase of the project, and that they improve transparency. The fact that nothing happens behind the door will certainly be saluted by donors.

Another proposed solution could be to organise two workshop sessions: one will be internal and will be useful for the staff, another one will be organised with donors.

A last remark concerned the fact that we should try when it is possible to organise workshops in the same time with other projects (potentially from other institutions) that involve the same stakeholders to reduce the costs.

2nd debate: staging a director of evaluation (DE) and a director de cabinet (DC)

The DC is complaining about the cost of maintaining an evaluation office. He argues that there are a lot of similarities with audit and does not see the point of keeping two units. It seems all the more true that audit evolved and now looks at results, performance and efficiency. Eventually, the DC argues that evaluation has no value added and mostly corresponds to the project findings.

For the DE, evaluation is different because it has learning drivers, proposes a broader picture and seeks the outcomes and impacts on a long term perspective. Evaluation is also much more qualitative. Furthermore, evaluation also works on mechanisms to ensure that its findings are used and useful for the organisation. On the other hand, audit is an internal function whereas evaluation is public and acts for transparency. The credibility of evaluation comes from independence.

Considering the costs, the DE also considers that he is working on a rational organisation, with a reliance on junior staff and Rome based agencies.

Concluding remarks and take-away messages

- Mixing national/international consultants when possible
- Support workshops (very important for ownership)
- Use M&E data
- Always think first to national consultants
- Think of value: how will you use the evaluation? The link value/costs should always be kept in mind
- Keep in mind the informal/formal dimensions of the evaluation
- About costs: it is probably much more efficient to do evaluations yourself considering that evaluations will be done by someone else otherwise, and the results will not be adapted to improve your organisation. Therefore, we should think about the costs of *not* having an evaluation
- The importance of restitution to the front line staff was underlined.
- It has been remarked that audit is indeed getting closer to evaluation as it is not as compliance based as it used to be. The risk of overlapping between audit and evaluation must be taken seriously.

Session 2.4 – **Evaluation and hindsight: assessing past interventions against yesterday's standards or today's wisdom and criteria**

Lead convener: **Fabrizio Felloni (IFAD)**

Collaborators: **Olivier Cossée (FAO), Natalia Acosta (UNDP), Simona Somma (IFAD)**

This session dealt with the fact that evaluators typically assess projects, programmes, policies and strategies that have been prepared in the past. However, some of the present standards, knowledge, criteria and practices might have not been available at the time of the design. Therefore, should evaluators assess only based upon standards, practices and knowledge available at the time of the design, or could they include our latest standards? Indeed, a strong case can be made in favour of the incorporation of standards and criteria considered as useful for stakeholders. In this situation, what role should the stakeholders play in deciding what criteria and standards (of the past and of the present) should be adopted for the evaluation?

Modality

The session consisted in a scenario simulation. The work was organized in three groups that discussed different scenarios or situation. The groups and questions were:

- Group 1: Changes in the evaluation methodology or in the objectives of the project during implementation
- Group 2: Changes in polices (e.g. Government or in the organization)
- Group 3: Changes in programme implementation context (e.g. War, conflict or natural disaster)

Questions:

1. What approach can your evaluation adopt to capture and assess changes?
2. How can you assess your organization's capacity to adapt to the changes?
3. What challenges would you find and how would you overcome them?
4. How can you ensure fairness in your evaluation particularly in assessing accountability?

Key messages

Approaches to capture and assess changes in evaluations

First, it is important to decide if the evaluation is still relevant and useful, given the changes in context. Some aspects to consider include: how big was the change, the extent to which it affected the programme, the timeframe of the intervention (as a determinant of the capacity to adapt to changes).

When there were changes in the programme objectives, it was noted that both the TORs and the evaluation report should acknowledge that there were changes and reflect on how the programme adapted to change/ responded. If the changes were formally approved, then the evaluator should consider those new changes in the evaluation design. If not, then he should give credit to what has been achieved but also indicate the mismatch. When assessing with the initial project objectives, evaluators should explain why this has been used. In this case, it might be useful to reduce indicators on adaptive management. However, there is a risk of reducing the utility or learning of the evaluation when evaluators do not take the changes into account – losing the potential to have an impact on the programme they were evaluating.

When there were changes in evaluation methodology or criteria: it would not be fair to hold the evaluand completely accountable if a new evaluation criteria is introduced. Instead, the evaluation could serve for learning purposes and could be used as an incentive for the project to include this criteria in the next phase. A useful tool is a theory of change; it can be used to map and understand the programme before the change and after the change.

It is also necessary to rebuild the results framework, based on the registered changes and considering the change in needs as they evolved.

Aspects to consider when assessing the organization's capacity to adapt

It is important to find the right balance between the capacity of country offices to respond to changes and the accountability (responsiveness).

Participants had mixed opinions. Some considered that if the programme was supposed to adapt, then the evaluation approach should be designed to integrate this aspect. If it was not supposed to adapt, then it would not be fair to hold the evaluand completely accountable. Other participants considered that even if the evaluation acknowledges the change, the evaluand should still be held accountable for their capacity to respond, consider risks or adapt.

It is important to reflect on how favourable policy change or negative policy change for the programme influenced the capacity to adapt and therefore the results of the programme.

If it is relevant to the new context, organisations should use responsiveness as an evaluation criteria (i.e. crisis or emergency evaluations)

Organisations should also work on the capacity to assess new needs; looking at the "paper trail": whether changes in implementation were recorded, can they be traced back?

Challenges and solutions

- Team needs to avoid having a "moving target" in their design and intend to set the basis
- Importance of watching for the level of engagement from stakeholders in approving targets
- Political sensitivity: it needs to be thought ahead (how to manage it in the evaluation)
- The fog of war: uncertainty.

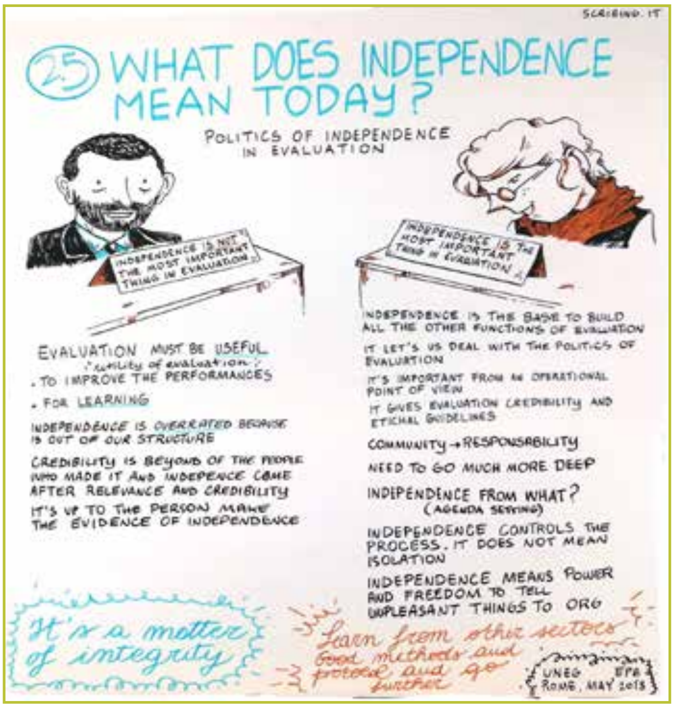
How to ensure fairness in assessing accountability of the programme managers

Participants stressed the importance of agreeing on any new criteria upfront in TORs (i.e. responsiveness), documenting the changes and express them clearly in the report, and "Playing nice" (understand that it is difficult to do well).

Concluding remarks and key take-aways

Evaluators should be transparent when integrating new evaluation criteria, fair and empathetic when conducting the assessment, and innovative when designing the evaluation and engaging with the evaluand and other stakeholders.

Session 2.5 – **What does independence mean today?**
Lead convener: **Jo Puri (GCF)**
Collaborators: **Indran Naidoo (UNDP), Chery Gray, Michael Spilsbury (UNEP)**



This session focused on the politics of independence in evaluations. The discussion provided an opportunity to re-assess some of the pitfalls of the evaluation profession, especially in the post-truth era. An important question that evaluation officers often face is: what does independence mean? The principle has been dealt with differently by various organizations. With the Green Climate Fund (GCF) setting up its new evaluation policy, it is a good time to reflect upon what could have been done differently in the interpretation and operationalization of the concept of 'independence'. Does independence imply a trade-off with 'use'?

Modality

The session will begin with an oxford-style debate about the importance of independence. It will then merge into a fish bowl discussion involving participants.

The statement of the debate is "Independence is not the most important thing in evaluation today". Two teams are invited to react on a partisan basis (agree or disagree with the statement). 5 min for each team.

Audience can participate after the first debate: a '6th chair' is available to take a non-partisan position; or participants can come at any moment and seat at the place of a debater to defend the latter's position.

Key messages

1st part: debate between the two teams

In round 1, teams exposed their arguments:

Team 1 - Position: independence is <u>not</u> the most important thing in evaluation today

There are very few truly independent offices, most of them are embedded into the organisations. For the few which have independent offices, it is legitimate to ask if the evaluation function is independent. In reality, it is thus observed that there is almost no independent evaluation function.

The most important is consequently to ensure that good work is done, which is what makes the office credible. Therefore, we should focus on building steering committees, boosting credibility by strong evaluation methods, expertise, etc.

Moreover, to be effective, the evaluation must ensure learning purposes, engage users and build the ownership of organisations. It cannot be achieved if evaluators are regarded with suspicion. The embedded function is interesting because it makes it easier to foster evaluation use when the evaluator is not regarded as an outsider.

Surely, the structural dependence can inspire fear. However, we operate on donors funds and donors really support evaluation. The issue is about political independence more than structural independence. The latter cannot be achieved in practice.

Team 2 - Position: independence <u>is</u> the most important thing in evaluation today

If we want to have a choice architecture to empower the evaluation profession, it has to start with independence.

Independence gives evaluators the choice architecture to design the evaluations we can possibly do.

Moreover, we must think about credibility, measurement. And we also think of independence as independence in analysis and conclusions.

What's more, it is difficult to provide a judgment if we are embedded in the organisation, it makes it hard to look up/ to have a wider perspective.

Surely, there can be a risk of isolation. The question is therefore to what extent can the evaluator make a judgment and be heard. It is important to manage to keep access to the organisation: go in and talk to people to have efficient evaluations.

However, at the end of the road, the most important is to be viewed as impartial.

As a matter of fact, it is observed that the best critical evaluations were done because evaluators could not be touched by the administration.

Concerning credibility, to be sure methodology is important, but in the end evaluators need space to talk to power.

In round 2, teams responded to each other arguments

Team 1 responding to team 2's argument

Surely, trying to become more independent is useful.

But we should first focus on making sure that evaluations can catch the attention of the managers. Being useful is more important than being independent. Consequently, independence is necessary but not sufficient and being credible and useful is the key.

If we get obsessed by independence, we will lose as regard to professional standards. We will be persuaded to be right, but only in our corner.

Team 2 responding to team 1's arguments

Yes, it is important to note that independence cannot and should not be reached at any time/level. A small structure that starts in evaluation should not move directly towards independence. It needs to be known for its credibility and methodology.

The idea is therefore to become so good that the evaluation office will be considered as useful, because its recommendations are taken seriously. Therefore, we must focus on the method and expertise because this gives the ability to become independent. Evaluation offices that are not professional, rigorous cannot pretend to independence.

If we are to be independent we also need to build frameworks to ensure that we will be useful. However, it must be noted that only independence gives us the ability to deliver rigorous evaluations.

2nd part: the audience was invited to intervene and share their arguments

Some participants stressed that, in the end, what matters is the issue of trust. There is an over-structuralisation of the debate. Structures should not monopolise our attention. In academic publications for instance, the structures of researchers are not so important: academics are trusted because of the quality of their paper, and that is all that matters.

However, presenting our profession as derived from science and methods is not necessary the best way of considering evaluation. Evaluation is about providing a judgment, and we must accept it.
That being said, it was emphasised that evaluators are not journalists. It is considered that the judgments must be informed, and that what led to that judgment must be explained.

But the idea that the institutional aspect was not the only one was relatively supported by participants. Indeed, even an independent structure can decide to make its life easy and avoid conflictual situations. On the other hand, in the end, even the people who finance embedded evaluations expect the evaluator to be independent. If evaluators say the same thing as the management, they are not considered useful.

Some participants affirmed that one cannot produce good evaluation if he is too dependent. Evaluators need a minimum level of set up (setting the agenda, etc.). Maybe those set ups are the most important.

Participants also supported the idea of a peer reviewing community, and a certain amount of standards, methods, etc. to ensure quality. But we need to go deeper than this if we think evaluation as a profession. For some participants, independence protects the profession and allows to control the process. However independence does not mean isolation. For some, it is up to the person to make the function credible.

Should evaluation office fight for independence? Yes. Should evaluation office be independent? Not necessary. There is a need to `build muscle` to pretend to be independent. If we had to focus on the most important, we should look at the head of evaluation and the head of organisation, their commitment the freedom of speech, etc. In the end, what really matters are behaviours in critical situations/wartimes. If those moments turn bad, it will undermine the credibility of the evaluation.

Moreover, independence is probably overrated. Government and donors do not really care about it. We can understand the protection offered by independence but it is not reasonable. For some, managing the trust deficit is the most important.

Independence is not important as the quality of the evaluation will make the difference. We must focus on having rigorous standards, professional credibility, quality, influence. Some participants argued that independence is only a mean to an end, and the end is to have impact, everything should be subordinated to that end, even independence.

However, several participants supported the idea that independence is most of the times needed to control extreme situations: independence matters at the margins. It was also signalled that the extreme does not come from management most of the time. Instead, it is governments which are scared of bad comments.

IFAD is evaluated as an implementing partner of a government. It makes IFAD very independent.

What is independence for them? It is about what is evaluated, who evaluates, who does the evaluator report to, and from whom comes the counterfactual. These are the most important questions.

From another perspective, participants argued that the problem is eventually to manage the conflicts of interest. Organisations are given money to change the lives of the people. They must prove that they do their best. Are organisations doing efficiently what they are committed to do? The idea is to follow that objective while minimising the conflict of interest.

Concluding remarks

Team 1 – Position: independence is not the most important thing in evaluation today

The architecture will not deliver independence. There many more challenges, we should not take everything back to independence. Independence is important but it`s one important things among others.

Team 2 – Position: independence is the most important thing in evaluation today

It is acknowledged that learning from other sectors, good methods, good protocols, etc. are determinant. Those things are untouchable. But we must go further.

It was remarked that when the mentimeter was used on the first day of UNEG to define the most important quality of an evaluator, very few answered `independence`.

Session 2.6 – **Full Disclosure? What evaluators don't write about**

Lead convener: **Olivier Cossée (FAO)**

Collaborators: **Fabrizio Felloni (IFAD), Michael Craft (OIOS), Marta Bruno (FAO)**

Evaluators like to think that they promote full transparency. In practice however, there are real life constraints to how transparent they can be. Those constraints go from the uncomfortable feeling coming from the evaluation of the performance of other non-UN partners to the fact that evaluators may be ill-equipped to analyse financial malpractice or individual performance. The aim of this session was to come up with an initial taxonomy of the most recurrent 'un-saids' in evaluations and also discuss and explore mitigation and lateral measures through which the information can be fed into the organizational learning and accountability mechanisms through other routes.

Modality

Three case studies were presented and discussed by participants in 3 groups, then each group presented its conclusions and key messages in plenary.

Key messages

Case 1: Fiduciary Issues - Facilitated by Fabrizio Felloni (IFAD)
The evaluation team unearths cases of sustained procurement fraud (or other instances where fiduciary issues come into play).

Is this even our job? Given the time and cost restraints under which evaluations operate, how does this potential fiduciary issue compromise the efficiency of the program/ intervention being evaluated – is this blatant instance just the tip of the iceberg – are there many others compromising value-for-money? Assess the relative size of the compromised programme/intervention within the program/intervention being evaluated. Does this impact of efficiency of the program, can it potentially erode the long-term sustainability of the intervention being evaluated?

Identification of Fiduciary Issues: Specific indicators serve as red flags drawing the attention of the evaluation team – higher than expected operating costs; high attrition rate in the project implementation unit; charging costs to vehicles: repetitive costs add up over the lifecycle of the project, a country-environment where corruption is rampant.

During the evaluation process: Establishing a tacit protocol on approaching the question of how fiduciary issues impact the intervention – will questions be raised tacitly? Will allusions be made? Is it acceptable to employ probing questions? To explicitly clear all misconceptions that members of the evaluation team are NOT members of the audit team.

Evaluation Report: The choice, and the manner in which fiduciary issues are articulated within the evaluation report fall within a sliding scale. While there are many cons to explicitly mentioning fiduciary issues in the evaluation report, one very strong benefit is that it requires a management response, ensuring that the issue will be revisited and the instance of fraud will be curtailed.

On the other hand, while not mentioning the issue in the evaluation report, or through other means such as reporting it internally, might seem like a safe choice, one also risks being perceived as thoroughly incompetent, especially when another evaluation/audit function follows the evaluation exercise, and finds recurrent, blatant instances of fraud, to their bewilderment, especially when the previous evaluation was relatively recent.

The participants shared a range of experiences on mentioning circumspect fiduciary issue in the final report: explicit mention of sustained corruption in the country which interfered with the successful implementation of the program, or procurement fraud by the project implementation unit; choosing to completely ignore the instance of corruption; and, alluding to instances of malpractice or procurement fraud which would possibly require further investigation by audit, while making the choice not to use strong language. The differences in experiences varied based on a range of factors: (1) whether it was acceptable to discuss corruption openly in that specific country context; (2) relationship with the government; (3) relationship with audit; (4) interactions with the project implementation unit; (5) alternate procedures for dealing with procurement fraud/other instances of malpractice in the specific country context.

What not to do:

1. Drawing unwarranted attention to instances of corruption/fraud through explicit mention in the report, if these explicit mentions are not substantiated with evidence.
2. Allowing external consultants to raise jarring questions on possible procurement fraud.
3. Not having hired consultants on the evaluation team in accordance with competitive procurement processes.

Possible alternatives:

1. Institute a written email trail with the project implementation unit bringing to their notice the reported instance of dereliction/fraud.
2. Coordinate/sequence interventions with the Audit function. Again institutions shared a range of institutional arrangements where audit reports were shared/not shared with the evaluation report allowing for flagging of possible fraud either because evaluation first unearthed it, allowing audit to follow, or vice versa: merits and demerits abound in each possible sequence. In cases where audit and evaluation conducted parallel operations, it might result in confusing stakeholders.

Conclusion: As premised in the book "Why Nations Fail", foul play is what holds many countries back in their path to development. However, the ability to bring it up, unearth relevant evidence as relevant to the subject matter being evaluated, and handle it is extremely context-specific on a) the country context in which the intervention is being evaluated, b) the individual organisation's culture, and c) the relationship of the evaluation office vis-à-vis the audit function.

Case 2: Quality of Leadership - Facilitated by Michael Craft (OIOS)

The evaluation unearths from interviews with various external stakeholders that the leadership of the program being evaluated is compromising programme outcomes.

What to do:

1. Understanding the scope of the problem
    - Understand that there are a spectrum of issues pertaining to questionable leadership: Getting too cosy with the current regime (understand the balance between political gain and recurrent trade-offs); Abuse of Power; and wrongdoing. In some other cases the dereliction of duty might not be as insidious: poor advocacy, for instance, but the impact on the programme might be even more hard-hitting. Hence, the degree of wrongdoing (moral charge) may be strikingly disproportional to the impact on the program itself.
    - Balancing of two equal and opposite interest: While seeking to ensure that no harm is done, there is an equal emphasis on ensuring due transparency in evaluating the intervention.
    - In some cases, the issues which surface during an evaluation can potentially be "Old" issues, or deal with the last person who was in power.
    - Impact of poor leadership can be profound on the program being evaluated: "Bad Management has concrete implications"
2. Evaluation approach
    - Determine the depth to which evidence is sought to be collected.
    - Strategic use of probing questions and triangulation to discern the impact of poor management on the program being implemented.
3. Evaluation Report
    - Ensure that the issue does not colour the remainder of the report, which can risk the entire exercise including utility.

What not to do:

1. Focus too much on individual performance, however central it may seem to the remainder of the evaluation.

Possible alternatives:

1. Informally report to the leader's supervisor or supervising office, e.g. send a letter while copying HR, to constitute an internal paper trail.
2. Differentiate between actual wrongdoing versus poor management, and respond appropriately:

report to relevant bodies in the case of the former; and seek other ways of analysing the problem at hand in the case of the latter, for example looking at reporting lines, roles and responsibilities.

Case 3: Political Factors - Facilitated by Olivier Cossee (FAO)

During an evaluation of electoral assistance in the post-conflict context of Complexstan, you come to realize that many complex political issues are intertwined with the electoral support you are tasked to evaluate, such as an election law that is still in parliament and could tip the scales one way or the other depending on who the law considers eligible to vote, numerous issues with past voter registration attempts, and delays in setting up the Electoral Commission. The Government considers that the process is going well and that its sovereignty should not be infringed by the UN. Donors paying for the elections are losing patience, complaining that they aren't seeing any progress. One of your team members argues that mentioning these issues in the report could hasten their resolution, but the Representative asks you not to mention this work in your report, as it is politically sensitive and there is a risk that speaking about it would render his task more difficult.

What to do:

1. Overarching principle: do no harm to the election process and its credibility, but instead thrive to strengthen its utility and credibility.
2. Evaluate on the basis of established evidence, in the most neutral fashion possible.
3. Focus on the role of the organisation you evaluate – e.g. if it is to support the election process in logistical and administrative terms, evaluate that role, *not* the whole election preparation process.
4. Involve, consult and inform widely.
5. Mention in your report the risks for the electoral support process to be affected by a number of political externalities -- you must mention the elephant in the room for the credibility and utility of your evaluation, because everybody knows it's there.

What not to do:

1. Don't get judgmental about the situation and/or the context; you don't need to encumber yourself with a political opinion on who would be the best president for Complextan, as this would render you unable to stay neutral in an already politically-charged environment.
2. Evaluation is a negotiation process, so don't start too low.
3. Don't use inflammatory language which could attract unnecessary, potentially damaging attention, and damage the peace building process.

Session 3.1 – **No baseline: how to measure impact?**
Lead convener: **Anna Henttinen (WFP)**
Collaborators: **Muge Dolun (UNIDO), Carlos Tarazona (FAO), Srilata Rao (OIOS)**

This session covered examples of different types of evaluations and situations where an organization has attempted to measure impact but has had no baseline to demonstrate change from before the implementation of the programme/ intervention/policy/strategy. Presenters shared examples of how they (their organizations) have creatively resolved this issue and worked to demonstrate impact, with reflection and discussion with participants. Both experimental and nonexperimental evaluations were covered through different practical examples.

Modality

The original modality is described below. However, due to logistical difficulties, the lead convener could not join this session, therefore the conveners had to re-organize it, and did not follow the original plan.

Safari / PowerPoint presentations

Participants were given several examples of different types of evaluations and situations where an organisation has had no baseline to demonstrate change from before the implementation of the programme/intervention/policy/ strategy. The situations are the following:

- When the development intervention targets multiple groups such as companies, policy makers, financial institutions, etc. (and not directly individuals)
- Where it is difficult to establish (statistically significant) control groups (i.e., due to project design targeting industrial sectors)

- Where measurement of performance against a benchmark might be possible (i.e., energy intensity of a sector) but doesn't allow to control for other contextual factors
- Where there might be ethical considerations in experimental/ quasi-experimental set-ups where deny intervention to a group might be difficult to justify
- Where the UN Agency has linked its performance to the achievement of SDG-related targets for which there is no baseline data available.

Key messages

No Baseline and field projects

Presenters shared possible approaches to address baseline data gaps when evaluating the impact of field projects.

This notably included the possibility of using a theoretical approach based on "Theory of No-Change" (Woerlen et al. 2016). The Theory of No Change (TONC) focuses on the roles of different stakeholders that are necessary to effectively implement projects. During the project the stakeholders face several barriers to fulfil their roles, e.g. lack of awareness, lack of expertise, lack of affordability. The TONC puts forward hypotheses regarding why certain causal linkages are in fact broken or why interventions cannot (yet) work in identified circumstances. The combination of the TOC and TONC lead to new approach for sketching the program logic.

They emphasised on defining impact through market transformation: looking for evidence of scale-up, mainstreaming and replication: how the programmes are lifting the barriers to energy efficiency and market transformation (policy, energy value chain, companies, financial sector).

They also suggested to take a different approach to evaluation or to use alternative evaluation techniques by using what information is available. 'Use creative solutions! Think outside of the box'. Also, remember that techniques work better when combined.

It is encouraged to ask other questions, use collaborative approaches and determine what to emphasize on to still be able to evaluate program/project.

There is also the possibility of using stacked programme design to provide intervention to individuals with temporal lags so to create their own control group. Besides, there is the importance of mixed methods: establishing counterfactuals and propensity score matching.

No baseline and non-field projects

In addition to field operations, several UN agencies have a normative and/or advocacy role. Measuring the impact of legislative frameworks, knowledge products and communication campaigns pose different challenges than measuring the impact of field projects.

Moreover, the progressive alignment of UN agencies' strategic frameworks to the Agenda 2030 requires appropriate evaluation tools to measure the extent to which UN Agencies are contributing to the SDGs.

Presenters shared possible approaches to address these challenges. Notably, UN agencies can align their results frameworks and evaluation plans to the Agenda 2030, as per the SDG guidance note developed by OIOS for the UN Secretariat. Evaluation units could also develop rubrics to measure progress in implementing programmes aiming at achieving SDGs; FAO shared an example of rubrics, which fed into a QCA exercise to compare programme performance at country level.

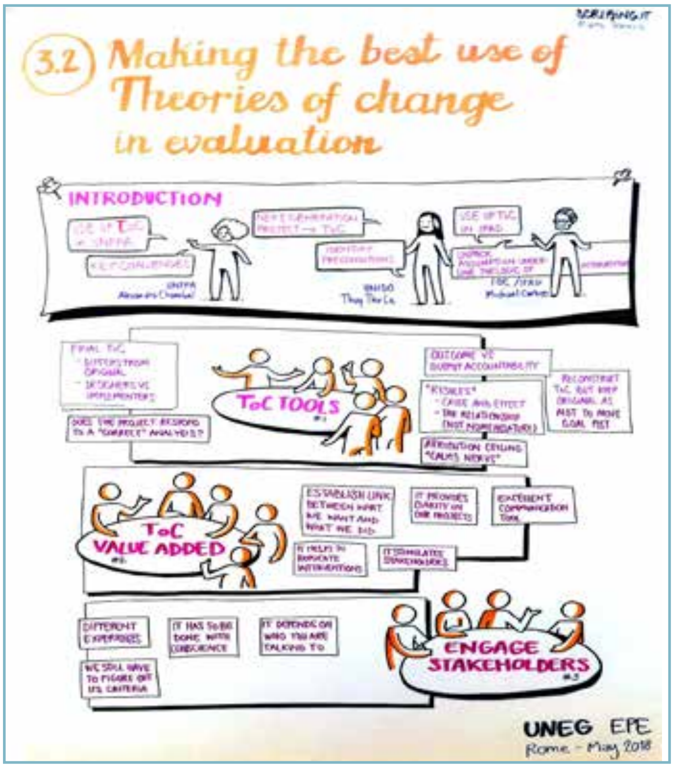
Concluding remarks

Evaluating the impact of field projects is not always feasible or necessary – there are much cheaper/ effective ways of getting equivalent results. Given the raising demand for robust evaluative data and the methodological challenges posed by the SDGs, It would be important for UNEG members to further develop experiences, identify best-fit practices and share knowledge to improve measuring the impact of UN agencies' operational, normative and advocacy work.

Session 3.2 – **Making the best use of Theories of Change in Evaluations**

Lead convener: **Michael Carbon (IOE/IFAD)**

Collaborators: **Alexandra Chambel (UNFPA), Thuy Le (UNIDO)**



The participants discussed the challenges and opportunities in using Theory of Change (ToC) in evaluations. This session started with a short introduction on what ToC is and how it is used in UNFPA, UNIDO and IFAD evaluations, followed by group discussions. Three group discussion topics were envisaged: i) what is the value added of using theory of change in an evaluation? ii) how do we engage evaluation stakeholders in the design and validation of a theory of change? iii) what are useful tools we can use to help us reconstruct a theory of change to support an evaluation?

Modality

Participants were divided into 3 groups. Each group had about 20 minutes to discuss each topic. The moderators and rapporteurs stayed with the same topic throughout the session but rotated from one group to another at the end of each discussion round. The moderators started of each round of discussion with a short presentation of what had been discussed on their topic in the previous round. This allowed all participants to contribute to all topics of the session.

At the end, the rapporteur for each topic summarised the key messages of his/her discussion topic.

Key messages

Topic 1. What is the value added of using theory of change in an evaluation?

ToC helps to assess impact or the progress towards impact at completion point: as it helps identify causal pathway from inputs towards impact. The key in ToC development is to identify pre-conditions (intermediate states or outcomes – the ´missing middle´) likely to bring about behavioural changes required to achieve the long-term impact, often referred to as system transformations which take time and rarely take place within the time span of a project.

ToC allows to assess the extent to which project activities correctly targeted the conditions necessary for impact to happen and to identify where projects stand on the impact pathways and what else should be done to reach impact. ToC is an excellent tool for assessing complex issues (programmes and themes). ToC originates from Complex Systems Thinking (Boundaries: Space and Time, Components: Domains & Agents, Interaction: different actors, Emergent properties – adaptive management) and helps to link interventions at macro, meso and micro level.

ToC makes the evaluation interesting both for stakeholders and evaluators. Evaluation criteria are for counting beans (accountability), ToC helps to look at strategic, long-term and bigger picture. ToC also allows discussion on technical content that stakeholders could refer to easily and more enthusiastically: good for drawing conclusions and lessons. ToC helps project managers and stakeholders and evaluators to better understand the processes that projects seek to influence. It helps replicating and up-scaling of interventions.

ToC adds rigour and credibility to an evaluation as it provides credible results and performance stories where M&E are incomplete or missing. But, evaluation teams need to study and analyse well the documentation before being able to reconstruct the ToC. ToC presents an integrated approach and better analysis of project context to guide key questions, consultation with key stakeholders, data collection and analysis, conclusions and recommendations. It also helps strengthening the assessment of all other evaluation criteria (relevance, efficiency, effectiveness and sustainability). Eventually, ToC also helps to identify key missing results to reach impact.

ToC is an excellent tool for communication and for monitoring progress and results during implementation. It tells stories about how and why short term activities & results contribute to long term changes and the conditions and assumptions that lead to change, in a simple way (1 page). ToC is good for presentation, interviews with stakeholders, media, evaluation report and knowledge products (evaluation brief, infographic…). It helps to explain to donors that the pathway to impacts and that impacts will happen beyond the project completion point. ToC helps to identify where the project stand in the impact pathway and what to be done to reach impact.

Topic 2. How do we engage evaluation stakeholders in the design and validation of a theory of change?

The engagement of stakeholders in the reconstruction and validation of ToC is not systematic across organisations. An important challenge is that project staff and other stakeholders are not familiar with the tool. Engaging stakeholders in this process is, however, very useful, as it helps to ensure that all parties involved in the evaluation share the same understanding of the programme strategy, reveals possible differences between what was planned and what was the reality, Identifies assumptions (external conditions that need to be in place for the ToC to work), and stimulates discussions on the programme logic and relevance.

The validation of ToC with stakeholders should be done in a careful manner taking into account the level of awareness about the tool and political sensitivities involved to avoid unnecessary tensions between stakeholders. Therefore, it is important to be strategic about who, when and how we engage in reconstructing and validating the ToC. Hence, there is a need to streamline the approach, which includes defining criteria to orient involvement of different stakeholders in different phases, depending on the subject matter of the evaluation.

Topic 3. What are the useful tools we can use to help us to reconstruct a theory of change to support an evaluation?

To reconstruct a ToC the following steps are generally followed: 1) identify change statements (outputs, specific objectives, goals etc.) from logical framework and design documents, 2) copy-paste those in text box shapes in Excel, 3) organise them according to the appropriate results level (outputs, outcomes, impact etc.) following the organisation's definitions, 4) connect them with arrows to indicate cause/effect relationships based on descriptions in design documents and logic, 5) add external factors affecting changes (see, for instance, risk statements in design documents) in block arrow shapes, and 6) fill any gaps in change statements, but use italics or colours to show that *you* have added these.

Participants made interesting suggestions and recommendations for reconstructing a ToC. They notably stresses the importance of not wasting too much time on definitions of results levels, but focusing on identifying the cause/effect relationships and be clear on the line where the project/organisation's accountability ends (called the "attribution ceiling" by one participant).

They also noted that the possibility to discuss the reconstructed ToC within the evaluation team and with the project team, and to use colour coding to show changes belonging to different causal pathways

They signalled the existence of a good software to design ToC diagrams, such as Excel, Visio, DoView etc., but recalled to take into account the access of the team members to the software and their ability to use it.

To reconstruct a ToC for a country programme, it is a good idea to start from the country strategy document. If this strategy does not exist or is not adequate, it can help to identify 4 or 5 quite generic outcomes that summarize well the majority of individual project outcomes. This represents the mid-level changes in the ToC. Project outputs can then be connected to the appropriate outcomes. The outcomes can then also be connected to the different impact domains of the organisation which are applicable to the country. Alternatively, different ToCs could be developed for key themes/components which are relevant and of interest to the stakeholders.

Participants also indicated a number of limitations of using ToC for evaluation. For instance, there is a risk that, when the evaluation team focuses too much on the ToC, it forgets that there may be alternative, and perhaps more efficient/effective pathways, to achieve the intended results. It is also important to consider that the ToC of a programme may have evolved in the course of programme implementation as changes are happening within and outside the ToC. Evaluator will either need to update the ToC regularly or opt for other evaluation tools which are appropriate for complex, rapidly evolving situations.

Session 3.3 – **Country Programme Evaluations – methodological challenges**

Lead convener: **Diego Fernandez (WFP)**

Collaborators: **Patricia Vidal Hurtado (ILO), Fumika Ouchi (UNDP), Natalia Acosta (UNDP), Mar Guinot (UNDP), Youri Bless (UNDP), Carlos Tarazona (FAO), Fumiko Nakai (IFAD), Simona Somma (IFAD), Ramona Desole (WFP)**

Each collaborating organization presented a CPE experience in which they faced methodological challenges, and then exposed the solutions they found to deal with those challenges. Discussion focused on priority CPE methodological challenges/issues, including over-dependence on qualitative dimension of mixed method, lack of ToC, engaging stakeholders, lack of baseline and time series data, contribution analytical method, implications of SDGs/UN Reforms of CPE methods and comparability of CPE methods.

Modality

The session aimed at sharing solutions for methodological key challenges identified in Country Program Evaluations (CPE). It consisted in 5-minute presentations from staff from FAO, ILO, WFP, IFAD, and UNDP around key challenges encountered and found solutions.

The audience then voted, using Mentimeter, to establish the 3 priorities that it wanted to discuss, based on the 6 main themes that were identified in the presentations from the different Staffs.

Afterwards, practice groups were created in the audience and each group discussed the challenges and possible solutions with colleagues.

Key messages

1st part: Presentation by the different conveners about the key challenges faced by organisations in CPE

FAO suggested two main discussions questions: what to do when we have a limited amount of time in a country to report on the results of the program and how to deal with lack of data availability? FAO referred to its experience dealing with a lack of theory of change (TOC) and baseline data in Kyrgyzstan CPE which assessed FAO's contribution to development results based on a country programme whose results framework had ambitious components.

IFAD finances long-term rural and agricultural development projects through the government. Country Strategy and Program Evaluations (CSPEs) conducted by the Independent Office of Evaluation of IFAD (IOE) cover not only the performance of such projects IFAD finances but also other dimensions and activities that are outside and beyond the project portfolio such as knowledge management, partnership building and in-country policy engagement. As such, stakeholders for CSPEs are diverse, ranging from the beneficiaries of development projects, i.e. the rural poor, their organizations, the government oficials and staff at different levels (national and local), NGOs to other donors, and such diversity does present challenges in meaningfull engaging stakeholders.

IFAD reports 4 key challenges in conducting CPEs:

1. Understanding the methodology of the evaluation and what we expect from the beneficiaries, especially from the portfolio level and the non-lending activities which scope is not always clear.

2. Budget and geographical constraints; IFAD prepares an assessment of a portfolio so they face limitations on the time that can be allocated for assessing a specific project; geographical spread is an important factor as they cover wide areas and sometimes it is difficult to reach beneficiaries;
3. There is also the issue of language for communication and interaction, since often the people in the field, beneficiaries or local staff, speak only local languages and require interpreters and;
4. Managing the expectations that people have from the evaluation and managing stakeholder interests.

ILO: CPEs at ILO are focused on country level results framework, which are called DWCPs. Before 2013 it was focused on a single DWCP. After 2013 CPEs started to focus on different DWCPs, clustered by theory of changes within sub-regions. Most CPEs are theory driven and goal-based and are supported by case studies, samples, interviews, surveys, focus groups and in-country visits. The main methodological challenges they face are: 1. Getting good theories of change (theory-based approach and goal-based setting); 2. Having a good link between an intervention and more than one outcome; 3. Outcome level monitoring data, and 4. Data analysis: sometimes, they have too much quantitative information from data analysis and sometimes qualitative information is scarce when they have low response rate to surveys.

When it comes to preparing the evaluation framework, they spend a lot of time on evaluability review. They also conduct synthesis reviews of past evaluations. They also invest a lot in the inception phase, for research and initial consultations with key stakeholders, to have enough time to readjust questions and methodologies and to reconstruct ToC and outcome level information. About low-response rate for surveys; they try to develop a stratified sample and they also make sure there are debriefing sessions in the country to validate findings.

UNDP launched its CPEs known as assessment of development results in 2002, to assess UNDP's performance in supporting the countries achieve their national priorities. Over the years, the Executive Board requested more evaluations done at the country level. Starting 2018, they would cover all countries submitting a new CPD for board approval. However, the key challenge has been doing this with limited resources, while maintaining quality. Last year, they conducted 8 CPE, this year is 15 and 2019 it will be 37.

In terms of solutions, they have taken several measures based on advisory panel and discussions:

1. Streamlining the evaluation process while strengthening evaluation methodology. For example, the number of missions has been reduced but they introduced more upfront data collection and analysis, e.g, rapid assessment before going to the mission. The idea is to use the mission to validate what they already know instead of getting new. They are also collaborating more with UNDP Audit team for joint missions and consultations.
2. Sharpening/narrowing the scope: e.g. i) reducing the number of cycles for review; ii) in some cases, focus on a few strategic issues; iii) making greater use of the past recommendations; and iv) shifting away from mechanical application of the 4 OECD DAC criteria to addressing three key evaluation questions, i.e. what exactly UNDP intended to achieve in a given cycle, to what extent it was achieved and what factors contributed to the results.
3. Introduction of a cluster approach, where multiple countries with similar development challenges are assessed individually but linked to their thematic evaluations. E.g. Implement 12 countries next year in Europe and the CIS region and checking also UNDP regional strategic approach
4. QA mechanisms, including strengthening of internal peer review mechanism, where all reports are now collectively reviewed by all evaluation managers to assess the construct of findings/ conclusions/ recommendations; how sensitive issues were dealt with; and how major questions in the ToR have been addressed.
5. They created more P2 and P3 positions in the office to augment the workforce.

They still have challenges, regarding ensuring quality in spite of quantity; and ongoing UN Reforms, with more focus on UN System Wide approach. The SG is repositioning the UN coordination of agencies in a country, under the same window. Question of exploring synergies of all participatory agencies under UNDAF.

WFP dealt with the issue of the lack of theory of change in the project documents so the TOC has to be reconstructed; staff turnover is also important in WFP. Use of corporate theory of change. We also discussed a challenge related to the dominance of qualitative analyses in mixed methods evaluations. A proposed solution was the use of secondary quantitative data to validate/replicate the original findings under evaluation.

The following challenges were identified:

- Challenge about reconstructing the TOC. The main challenge is the lack of institutional memory due to the high staff turnover.
- Outcome Data unavailability and low quality (fragmented; not consistently collected over time; no baseline), In the case of emergency operations, which last less than one year, do not require collection of outcome data .
- Moreover, there is often a bias towards the present. It is hard to find protagonists from the portfolio years. Considering also the time lag between the closure of some of the programmes, most of the time there is limited institutional memory resulting from the high staff turnover, within WFP and its co-operating partners and national government

Solutions identified:

- The main proposed solution to the lack of ToC was to devise lines of enquiry that cut across similar WFP activities in different programmes/operations and other cross cutting issues so that they are not treated in a compartmentalised way.
- Monitoring data has been complemented with the data available at national level; Secondary outcome data have been augmented with info collected by other partners, UN agencies as appropriate, and triangulated evidence with other sources;
- Telephone/skype interviews have been conducted with key staff who had left the country but used to be based there for the first early years; surveys addressed to WFP staff and partners have been also conducted.

Later, people voted using Menti: the topics with most votes were: 1) Availability of data, 2) Balancing Quality and Quantity and 3) Use of Theory of Change.

2nd part: Audience was then split in different groups, the following paragraphs report the main discussion points

TOC Change Group:

The group reported the typical issue of the lack of TOC in project documents and the fragmentation of activities to evaluate during a CPE.

Solutions: the level of fragmentation across agencies is very different therefore, they could not come up with only one solution but they suggest trying to reconstruct the existing TOC building on log frames and strategic plans and to show some flexibility to adapt the TOC that moves along in the evaluation. Other solution: develop separate TOC by theme.

Stakeholders engagement

The involvement of stakeholders at all levels (from government to beneficiaries) in the evaluation process is crucial time and budget constraints in conducting evaluations represent the main challenges to stakeholders involvement. Moreover, evaluations are conducted in remote areas where often only local dialects are spoken. One way to address this issue is by hiring local consultants able to speak the language of the beneficiaries and ensure the diversity of the team. The group also suggested to collect more data at the beginning of the mission and to clearly communicating key milestones with stakeholders including country offices. Last solution to evaluation fatigue: joint missions when evaluations are carried out at the same time in a country.

Maintaining quality at scale

Clustering approach between different organizations was suggested as one solution as well as filling questionnaires before the mission so that data can be validated more easily during the mission. Another solution: putting a roster of regional consultants or thematic experts.

Cross Agency Collaboration in CPEs

UNDAF and individual CPEs. Solution to have more coordination between CPEs by loading work plans on UNEG website and exchanging on evaluation content. Optimal solution: to have a system in which all evaluation offices communicate. Also, at the central level talking to the UN Country Coordinator and explaining the benefits of a joint evaluation exercise; piloting this option in one country was suggested.

Balancing Qualitative and Quantitative in mixed-methods

Issue of balancing types of data. Solution: working with management of the organization; trying to get as much information as possible from stakeholders, including national level quantitative data; establishing minimum requirements for data typology in evaluations. Issue of costs: solution could be to reinforce M&E systems of the organization and working better with national counterparts.

Availability of Data

First, defining what availability of data is; how to collect the right data on activities and outcome levels. Optimal scenario was discussed, including collection of diverse socio-economic and geographic data.
Solution: doing an evaluability assessment to know what is available and use proxy indicators and data from other sources in case of unavailable data. There are still barriers in sharing data among agencies; agencies should collaborate on this aspect.

Concluding remarks

The possibility to continue the discussions by creating a Community of Practice with the focal points of each UN Agency was presented at the end of the session. Emails of the focal points were shared:

Topic	Lead Agency	Topic Focal Point
Balancing QUAL and QUANT in mixed-methods	ILO	Patricia Vidal Hurtado vidalhurtado@ilo.org
Use of ToC	WFP	Diego Fernandez diego.fernandez@wfp.org
Engaging stakeholders	IFAD	Simona Somma: s.somma@ifad.org
Availability of data	FAO	Carlos Tarazona Carlos. Tarazona@fao.org
Cross agency collaboration in CPEs	UNDP	Fumika Ouchi fumika.ouchi@undp.org
Maintaining quality at scale	UNDP	Natalia Acosta natalia.acosta@undp.org

Session 3.4 – ICT for data collection and analysis in evaluations

Lead convener: **Katinka Koke (UNITAR)**

Collaborators: **Shravanti Reddy (UNWOMEN), Juha Uitto (GEF), Hansdeep Khaira (IFAD), Brook Boyer (UNITAR), Olivier Vandamme (UNITAR), Michael Craft (OIOS), Luisa Belli (FAO)**

The discussion topics focused on how different ICT tools such as machine learning, geospatial data and big data can be used in evaluations for both data collection and analysis. Participants could learn from four out of the six different presentations organized by the collaborators that talk about their organizations' experience. Participants then had the opportunity to interact with the collaborators by asking questions or sharing their own experiences, facilitated by the collaborator's discussion question.

Modality

The session discussed how can ICT support and improve data collection and analysis for evaluations, due to the increasing potential ICT tools provide for high quality evaluations. The topic was aimed to keep evaluators abreast on cutting edge innovation that is complementary to evaluation practices and can be seen as a follow-up to the international conference organized by IFAD in summer 2017 on the same issue.

The session showcased past and current uses of ICT tools in different UNEG agencies as well as discussed the future potential for using these tools. The goal was to have participants better understand ICT tools such as GIS data, machine learning, and big data.

The structure was a world café "show and tell" of four rounds, each 15 minutes, where participants move around the room listening to the collaborator's presentation of their topic followed by a short and interactive discussion.

1. Short introduction by UNITAR
2. One minute elevator speeches from collaborators to introduce their topics
3. Four 15-minute rounds of presentations and discussions
4. One-minute wrap-up of take away messages from collaborators and lead convener

Key messages

GEF

GEF presented the institution's lessons learned from using geospatial data and machine learning in the evaluation of natural resource interventions with a mixed methods context. They discussed two examples on forest cover loss and land degradation that used ICT tools to address evaluation challenges.

Discussion question: What are the lessons learned from the application of geospatial data and machine learning to address complex evaluation challenges such as lack of baseline, sampling bias, counterfactual selection, identifying confounding factors and how can these methods complement traditional evaluation?

Some of the challenges and limitations highlighted were:

- High computer power and technical skills required;
- Uneven availability and accuracy of contextual variables;
- Difficultly to answer "how" and "why" questions;
- Need to conduct field verification and "groundtruthing" as well.

The collaborators' main **take-away messages** were that geospatial tools can be very powerful to quantify environmental challenges and to reconstruct counterfactual scenarios. However, to use GIS tools, the starting point is geo-referencing project locations. Often, these are not available in project documents, thus it needs to be done. Geolocating project sites is the most intensive part of the whole effort, both in terms of time and costs. Also, using these types of tools requires high skills capacity in evaluation teams; but University collaboration can be an opportunity to use talented data analysts.

IFAD

IFAD's presenters discussed their report on the IFAD-supported "Agricultural Support Project in Georgia". They presented how the results of an impact evaluation can have additional validation using geospatial analysis from a statistical perspective, and explained how the before/after control/impact (BACI) contrast measures the intended outcome. The study focused on the change in vegetation cover in Georgia from the rehabilitation of irrigation canals. The challenge faced was to select a control group without a valid baseline, so they used time series satellite imagery with a BACI contrast statistics (similar to the difference-in-difference method). Results of the measured outcome (change in irrigation use) were close to those obtained from using a household survey.

Discussion question: How can challenges in ex-post impact evaluations of an intervention (lack of baseline treatment and control) be improved with the validation of the geospatial results?

Some take-away messages were that the use of GIS data increased the statistical vigour of the impact evaluation, helping to overcome issues in baseline. These are freely available, open sourced data that all evaluators can have access to. These methods require good M&E systems to collect data and use of technology such as smartphones. Although, the scope of these ICT tools is not applicable to all situations, they can be very useful for measuring outcomes of certain interventions. Nevertheless it is not recommended to use these tools as a standalone method for evaluating agricultural projects but in tandem with some ''ground-truthing."

UNITAR

UNITAR presented the GIS maps that were used for monitoring and evaluation of projects with UNDP and BADEA, with both high and low resolution of the satellite images. The maps focused on infrastructure such as construction of roads, schools and hospitals. UNOSAT's maps are useful in areas such as humanitarian affairs, development, environment, human rights and others. UNOSAT is not only offering GIS services to other UN organizations and partners, but also providing training on analysing GIS images, which could be of interest for evaluators amongst others.

Some of the opportunities with GIS data concern accessing an archive of historical images but also real-time imagery. It is particularly useful when physical access to regions or countries is limited, e.g. due to security reasons, difficult logistics, or high costs. GIS data is not very costly in comparison to traditional data collection tools e.g. in the framework of missions. It is important to note that while it does not replace traditional data collection methods, it serves more as a complement to other data collection tools throughout preparation and planning of missions or interventions.

A few limitations are the scope of imagery for outdoor space, e.g. no indoor images. Also, small objects are not visible. Weather conditions can influence the quality of the images too. Furthermore, it is impossible to evaluate whether a project functions properly but only that it was built, e.g. a water well may have been built, but is there actually water?

Discussion question: What are the advantages and challenges faced with using GIS mapping and data for evaluation?

The take-away messages were that GIS data is useful for quantitative analysis, however it is less useful for qualitative analysis. Also, GIS always complements evaluation methods rather than replacing traditional methods. Nevertheless, there is a great opportunity to overcome challenges such as "too-big-to-measure" issues. However, it is more difficult to answer questions on how people are actually affected by the intervention.

OIOS

OIOS presented recent experience using social media analytics to capture outreach, media coverage and perceptions of the UN using the ICT tool Crimson Hexagon. This application was used in two evaluations conducted by the Office of International Oversight Services (OIOS) Inspection and Evaluation Division (IED) in 2017-18[6] in partnership with the *United Nations Global Pulse* initiative of the Secretary-General.

Given the increase in the volume/velocity/variety of available data, evaluators have at their disposal large amounts of information for potential analysis. Social media data is one area where evaluators can for example collect, code and analyse publicly available information in order to map conversation around selected themes or social media accounts.

Discussion question: How could this type of ICT instrument fit into the "evaluator's toolbox" for data analysis, and how can various forms of big data be used responsibly to meet evaluation standards?

The take-away messages are that these tools are costly but potentially efficient and useful for assessing specific intervention types. Participants pointed out during the discussion that this kind of tool could be used to gather data around the effectiveness of UN work related to advocacy, behaviour change, and influencing social norms. The tool can also help establish a contextual baseline for understanding conversation around selected topics on social media platforms. Participants equally noted the importance of taking into account major limitations of social media in representing only segments of the population, and therefore introducing strong bias in the results.

FAO

FAO discussed its "Presentation of new cloud geospatial technology in support of evaluation – collaboration between FAO and Google". FAO in partnership with Google has recently developed some tolls that enable immediate access and analysis of geo-spatial information related to environment, land use and climate variables. These tools have already supported the formulation and the implementation of several projects in different fields and are being piloted in evaluation of FAO land related activities. This presentation showed some of the applications/use of the tools for evaluation.

The tools presented by FAO/OED were:

- Collect Earth, a free and open source software for land monitoring developed by FAO. Built on Google desktop and cloud computing technologies, Collect Earth facilitates access to multiple freely available archives of satellite imagery, including archives with very high spatial resolution imagery (Google Earth, Bing Maps) and those with very high temporal resolution imagery (e.g., Google Earth Engine, Google Earth Engine Code Editor).

6 Evaluation of the United Nations Department for Public Information; and Evaluation of the United Nations Offices of the Special Representative Secretary-General for Children and Armed Conflict, Sexual Violence in Conflict, and Violence against Children.

- Earth Map, a FAO-Google tool for quick historical analysis of environmental and climate parameters, where you can catalogue your specific area of interest.

Discussion question: How to integrate geo-spatial information in evaluation planning design/impact evaluations?

The take-away messages focused on how the tools are user-friendly, open sourced and free to complement statistics. The analysis can be done in a few seconds for "longer periods" of time. The tool is less costly or time consuming than other traditional evaluation tools.

FAO's conveyors also stressed the importance of having geo-referenced data at the beginning of the project. Otherwise it would be difficult to evaluate the project later on without the exact spot or location of the GIS image to understand in the field what happened. FAO stressed the application of ICT tools in areas such as landscape changes, which can change the way evaluators do evaluation with a new dataset available (applications: rehabilitation of canal; open areas; map prolonged droughts for Climate Change; land use change can be captured; tracking natural disasters). This can replace the need to do extensive background studies with visual data. Collect Earth and Earth Map are user-friendly tools that have free tutorials available that may provide more access to data for evaluators.

Concluding remarks

In order to continue the conversation, Katinka collected names and emails to create a UNEG Slack group, in addition to joint efforts with IFAD to revive the existing community of practice on ICT for Evaluation initiated by IFAD in 2017. Here are also the names of the collaborators as well as their emails for further exchanges and possible future collaborations.

Name	Organization	Email
Juha Uitto	GEF	juitto@thegef.org
Carlo Carugi	GEF	ccarugi@thegef.org
Hansdeep Khaira	IFAD	h.khaira@ifad.org
Giancarlo Pini	WFP	giancarlo.pini@wfp.org
Olivier Vandamme	UNITAR	olivier.vandamme@unitar.org
Brook Boyer	UNITAR	brook.boyer@unitar.org
Michael Craft	OIOS	michael.craft@un.org
Luisa Belli	FAO	Luisa.Belli@fao.org
Marcelo Rezende	FAO	Marcelo.Rezende@fao.org
Danilo Mollicone	FAO	danilo.mollicone@fao.org

Participants who expressed interest in a community of practice:

Name	Organization	Email
Lavinia Monforte	FAO	Lavinia.monforti@fao.org
Albert Garcia Hernandez	WFP	Alberto.garciahernandez@wfp.org
Jane	UNICEF	jmwangi@unicef.org
Gillijus	IFRC	Josse.gillijus@ifrc.org
Johannes Federkeil	IFAD	j.federkeil@ifad.org
Fortunato Lambiase	SENATO	Fortunato.Lambiase@senato.it
Arianna Spacca	WFP	Arianna.spacca@wfp.org
Srilata Rao	OIOS IED	rao3@un.org
Julie Thaulougan	WFP	Julie.thaulougan@wfp.org
Anne-Claire Luzota	WHO	luzota@who.int
Fumika Ouchi	UNDP	Fumika.ouchi@undp.org
Ana C. Aza	FAO	anaclaudiaaza@hotmail.com
Javier Guarnizo	UNIDO	j.guarnizo@unido.org
Sarah Orenbo	OCHA	orenbo@un.org

Session 3.5 – **Evaluating policy support**

Lead convener: **Veridiana Mansour (FAO)**

Collaborators: **Olivier Cossée (FAO), Alena Lappo (FAO), Juha Uitto (GEF), Carlo Carugi (GEF), Arild Hauge (UNDP), Sasha Jahic (UNDP), Andrew Fyfe (UNCDF), Felix Herzog (UNESCWA)**

This session aimed to explore the contemporary complexities and challenges of evaluating policy support through shared experiences of the different methods and approaches used by evaluators. The five stages of evaluation (agenda-setting, policy-setting, implementation, M&E/data, policy environment) were discussed in five working groups. At the end of the session participants were expected to have: i) a common understanding on the challenges and complexities involved in this type of evaluation; ii) knowledge of different practices experienced by the different organizations; and iii) a set of methods to evaluate policy support.

Modality

Introduction to the session, and brief presentation on policy support in order to explain why the discussion was relevant and timely. It dealt with SDG context, complexities of the policy process, type of policy support, goal of the policy support, long term policymaking process versus short term interventions

Then, each collaborator had four minutes to present their experiences, and each presentation addressed the following topics:

- Indication of the stage of the policy process to which the support was provided (i.e. policy setting; policy implementation; M&E/data; policy convergence);
- Brief overview of the evaluation, project or tool, including the type of support that was evaluated;
- Challenges and lessons learned;
- Approach and methods applied.

They focused on at least one stage of the policy process in order to explore – and demonstrate to participants - the variety of challenges and complexities comprised in this type of evaluation.

After the presentations, participants were divided into five groups based on their choices. Each group represented one stage of the policy process (i.e. agenda-setting; policy-setting; policy implementation; M&E; policy convergence).

Groups hold discussions on the main challenges and complexities of evaluating policy support as well as on ways to overcome them. During the discussions, one facilitator recorded the key challenges and solutions in a flip chart. In parallel, a data visualization specialist illustrated elements of the discussion.

Each group had 2 minutes to present the results of their discussions. At the end, the flip charts representing the different steps of the policy process provided participants with an overview of the key challenges and solutions to evaluate policy support. At the end of the session, the creation of a working group on the topic was proposed to better contribute to the 2030 Agenda for Sustainable Development.

Key messages

Challenges:

On the 'WHAT' - Defining policy support

Challenges concern the lack of clarity on what is support to policy-making, and how it can be distinguished within the policy process. It also concerned the difficulties in expanding the results framework to include the 'policy for what' question as in the case of interventions supporting market development, including policy actors.

On the 'WHERE' - Understanding the context

Challenges concern the understanding of the evolving policy context and actors involved in particular stage of the policy process, and the existence of contradicting policies and support provided by different actors related to the same subject area.

On the 'HOW' - Using available data & methods

Challenges concern the availability and reliability of national data or the establishment of M&E systems within the changing context/policy environment.

Evaluators must also be careful to understanding the evaluability of specific aspects of the policy support; being realistic about the expected outcome of the support and the feasibility of evaluating it (e.g. realistic Theory of Change)

Concerning the contribution and attributions tracking, it was noted that the more actors and variables targeting/ involved in the same outcome, the harder is to assess contribution and attribution of the supported provided

It is critical to measure the impacts because results and variables are often not tangible; hard to identify counter-factual.

Possible solutions

To better define policy support

Organisations should focus on identifying the type and goal of the policy support as well as the stage of the policy process to which the support is provided. They should also tailor the analysis to the type of policy support evaluated.

Participants also invited to draw a line of accountability for the policy support evaluated (e.g. there are examples when it is not possible to attribute particular contributions/results to a particular event/project/action); and to specify carefully the results framework so that both direct and indirect results can be captured using appropriate mixed-methods data collection approaches (cf. work on this done by Consultative Group to Assist the Poor supporting measurement of improved financial inclusion systems).

To better understand the context

It is interesting to involve evaluators who are not only subject experts but also have a solid knowledge of political economy within particular country. A political economy analysis and policy influence mapping (key players, key influencers, observed and unobserved variables, other complementary and contradicting policies, etc.) can also be conducted.

Eventually, evaluators must be careful to tailor evaluation approach to the country context and design precise evaluation questions.

To use available data & methods

The importance of conducting in-depth evaluability assessments was firstly stressed. Evaluators should also be careful to have a clear and realistic evaluation ToC, which should be flexible in changing approaches and assumptions as a response to the changing policy context, focus on the intermediate states of change, and distinguish between policy support outcomes and outputs

The necessity of establishing clear and pragmatic indicators was also signalled, with the importance of tracking the process for outputs and outcomes (e.g. contribution analysis; review of old and new policy documents; multi-causality comparisons; mapping contribution through extensive desk review and stakeholder consultations; prioritize mixed methods).

Concluding remarks

- Although the discussion was very rich, the uneven knowledge among participants as well as time constraints did not allow more in-depth discussion.
- There was an overall consensus about the relevance of the topic as well as on the importance of continuing this discussion to create a useful material/framework for evaluating policy support.

Session 3.6 – Emerging principles and lessons learned for the development of organizational evaluation policies
Lead convener: Jo Puri (GCF)
Collaborators: David Ridersmith, Susanne Frueh (UNESCO), Andrea Cook, (WFP), Ludovica Soderini (GCF Board Member), Masahiro Igarashi (FAO)

This session consisted of an interactive discussion on what we are learning from evaluation policies around the multilateral world and the extent to which these are helpful or act as bottlenecks in evaluations and reporting and measuring effectiveness, efficiency, sustainability and relevance. Discussion also included an inquiry into where donors feel evaluations should reside in the reporting for donor agencies. A group of agency representatives contributed to the discussion both with each other and the audience.

Modality

After an introduction, each panellist talked about three aspects related to the design of evaluation policy: 1) Credibility and Measurement; 2) Usefulness and 3) Impartiality and Independence. Each panellist was asked to speak for no more than 4 minutes as a starting statement.

At the end of each round, the audience inter-acted with the panel and provided comments. There was also participation from the audience through Menti website, who were asked three questions on Menti:

1. In one word please describe what you think is the biggest challenge to credibility in evaluations?
2. In five words or less, please describe methods that you have found effective in increasing the use of evaluations
3. Evaluation functions should be independent: Please say 'Yes' or 'No'.

Key messages

1st part – Question to the panel of collaborators: to what extent have evaluation policies, in your experience, helped to foster credibility and measure results?

UNESCO

UNESCO proceeded by referring to 3 policies:

The 2006 UNDP Evaluation Policy: In 2006, UNDP did not have an Evaluation Policy. The norms and standards in evaluation were developed at that time. They looked at criteria for credibility, independence, impartiality, quality, timeliness, and provided a guide in UNDP around evaluation functions. In 2010, there was an update of the 2006 UNDP Evaluation Policies with additional support to credibility, and a real framework for Evaluation Policy was created. The weakness of the framework was related to the fact that there was no strong statement around decentralized evaluations.

Uganda Evaluation Policy 2011: this policy was developed on Monitoring and Evaluation, in an internal evaluation unit within the Prime Minister Office. It helped supporting the development of an Evaluation Government facility and outlined very specific requirements for measuring performance, including budget monitoring. However, although the Policy has since been applied at Central level, in many parts of Uganda it has not made a real difference.

DFID: 2011 the UK Decentralized the evaluation function from its office to the Programs and they set up an Independent Commission to embed evaluation into programs. However, one shortcoming of the Policy is related to Quality of evaluations at decentralized level.

FAO

Evaluation policy ensures credibility, supported by the quality of the evaluation, its methodology, independence and impartiality of the process. One of the underlying principles is related to *accountability* and articulates around various questions: who is accountable for evaluation reports and has the final word; who will really insure quality and that evaluations are evidence -based and useful.

Decentralized evaluations should be high quality evaluations because of credibility; in many cases, decentralizing evaluations includes independent external evaluators. Quality insurance mechanisms should be established in this respect, not only for post-facto checklists but also to ensure that methodologies and protocols are followed and that data can be provided and checked.

Evaluability should be the responsibility of the Evaluation Manager and results monitoring is the responsibility of all the team. In the case of a mid-term review it is manager's prerogative to check processes but in the case of a mid-term evaluation there is a need to develop quality insurance processes.

UNESCO

UNESCO Evaluation policy has contributed to the credibility of the function; as FAO said credibility is a very complex component of many aspects, including reputation. Therefore, UNESCO has put credibility as a combined function, with the investigation function and, sometimes, they may suffer from an issue happening with the investigation side. The highest risk comes from decentralized evaluations, as sometimes it is hard to control everything that is happening in the field. A lot of work needs to be done so that credibility remains for decentralized evaluations.

Contribution to results: strategic results reporting should be developed with evaluations designed against identified knowledge gaps.

When reports come from the field, they do not have to sign them off; they try to do a meta-evaluation of the results; they provide backstopping but this is decentralized.

GCF

From a Board Member perspective, they need to show what is effective and to have a credible evaluation. Management wants to see results, especially as a donor country, and wants evaluations that are independent, using solid methodologies.

Another aspect of credibility: the professional evaluators should have experience and be able to communicate in the right way with people with different perspectives, cultures and backgrounds.

2nd part – The floor was then opened to interactions with audience, initiated through Mentimeter

Question 1: What do you think in one word has threatened credibility of evaluations?

The audience used Mentimenter to answer. The most reported answers were: Poor Quality, Bias; Resources.

OIOS: this institution performs systematic reviews. They try to use the UNEG standards to assess the evaluation policies of the UN Agencies but they feel that UNEG Standards are sometimes daunting, since some UN Agencies do not really have so much capacity when it comes to developing Evaluation Policies. They advised a lighter version of UNEG policies could be developed.

UNICEF: evaluation policies really reflect the organization culture. For instance, UNICEF is very decentralized and so is the evaluation function in this institution. One of their lessons learnt about ensuring quality at decentralized level is that quality assurance should be included and reported right from the Terms of Reference. They have developed an External Quality Assurance system, at TOR; Inception Report and Draft Report stages that is mandatory for all evaluations, which ensures quality and consistency. This approach led to a majority of high quality evaluations last year.

UNFPA: many questions were raised:

- How far is the evaluation policy credible in your organization? Is an evaluation policy credible if an organization decides to merge the independent evaluation office with other functions, such as the investigation function?
- Is the evaluation policy designed to measure performance of the evaluation function, against KPIs for instance?
- About quality: should evaluation policy be a little more elaborate in defining quality, against the UNEG standards for instance?
- About the funding: the target for the evaluation function was set a 3%; where does this 3% come from? Was there any prepared financial analysis? It seems it was designed to protect and enhance funding for the evaluation function but it was not based on any financial analysis.

FAO: evaluation is only a tiny piece of what happens in our organizations. Sometimes evaluation policies were developed as a result of processes that created a momentum to reflect on evaluation which can come from external factors or built from within. Evaluation Policies need to be practical and integrated into another, higher-level, policy context.

UNIDO: evaluation policies seem to be violated in various organizations when the evaluation function was merged with other functions: how can we have credibility of the evaluation policy if it is not protected and violated?

UNESCO: FAO's point of momentum is important. It is an opportunity to speak about the evaluation function per se, including the decentralized evaluations. At UNESCO, they managed to flag missed opportunities related to decentralized evaluations which helped discussions with management and the member countries. Two processes took place in parallel: communicating about evaluation with the partners and decentralizing the process of evaluation. They have a strategy to implement the policy, such as online training.

Issue of funding: the UN systems practice is 3% for

Evaluation, which is not a lot at UNESCO. They decided that the 3% of budget for the operational funding helped them having a good estimate to plan evaluations. The 3% should also be at project level to have more M&E staff inside the project teams. There is a need to strengthen M&E and building data capacity within projects. The Evaluation policy helps giving credibility to the process, and to bring data to the table to show the need for evaluation.

How can we protect evaluation policy? They have a solid advisory committee that audit the functions they have. They have evaluation competencies along with auditors, which brings balance on the need for evaluation, the need to protect independence. They recommend having a committee, with the right people, that is an interface with the Member States and Management. There is no ultimate defence but a need to develop strategies to defend Evaluation Policy and functions.

GCF about the budget allocated to Evaluation Policy: there is a need to have more accurate estimate and that the numbers are explained to all stakeholders.

FAO: what we mean by quality and quality control? Evidence based is the most important aspect that is why there is a need to keep primary data available on demand.

UNESCO: the culture of the organization will determine which type of evaluation policy can be prepared and effective. Regarding the option of including evaluation into program budgets; this was part of DFID Evaluation Policy proposal. Then about UNEG standards: they were designed to be a basis, but not an absolute standard so they could be applied differently.

Question 2 what, in five words, or less, has helped foster use of evaluations, in your view?

Mentimeter's answer from the audience: Participation, Relevance, Credibility, Quality.

GCF on the usefulness of evaluations: they need to be independent, well discussed with the management and all stakeholders involved in the discussions. Data should be evidence-based to enhance credibility. The evaluation should not be negative, in the sense that it needs to suggest solutions to the identified and reported problems.

UNESCO: the participation of key stakeholders is very important, such as having a reference group, it is a heavy investment in the process but it pays off in the end. The same could be considered with Management responses: it is important to keep the dialogue open, as well as with member states (in spite of possible political pressure). Communication is the key.

FAO: there is a need to have a real purpose of the evaluation, which is that the Organization works better to fulfil its mandate. For each evaluation, the usefulness of the evaluation should be reported and explained. Stakeholder buy-in and engagement is also instrumental; UNEG has defined the principles related to UNEG engagement with stakeholders.

UNESCO: In Uganda, before developing performance reviews they were identifying issues that could be picked up by the government. By picking a topic that came up in the political debate there is a huge amount of interest in the evaluation from counterparts.

Finally, independence and impartiality: sharing best practices, regarding the interaction between impartiality and the use of evaluations. Comments from the audience:

Global Fund argued that they have impartiality, in the structure of the Evaluation Unit, within the fund, and they are directly reporting to the Board.

According to UNFPA: relevance is key to the use of evaluations, as well as engagement throughout the evaluation process; this is where the right balance is hard to find, between engaging with stakeholders and maintaining independence. In their case, they try to find a balance between summative evaluations that look more like accountability, and developmental evaluations. They also develop participatory workshops where external evaluation teams come and try to facilitate dialogue with Management. They can do it thanks to the structure, as they directly report to the Board.

For OIOS, engagement is not losing independence but coming up with a topic that interests and resonates to the stakeholders.

Question 3: in your view, is independence of the Evaluation function important?

Mentimeter's reported answers: 25 said Yes and 4 no.

Question to the panellists: what would you do different now if you had to do an Evaluation Policy?

FAO: building on UNFPA intervention; on independence, we have institutional and behavioural independence, but sometimes this is not made into the policy and then creates confusion from Management. On budget allocated to Evaluation, it should be increased. On decentralized evaluations: there is a need to have a stronger evaluation function or practice in the field.

UNESCO: the culture of the organization is important and evolves so it has to be monitored closely. They have prepared a baseline to try to report on what is working. Funding issue is relevant as there is some push back on the 3%, so they could have explained this aspect better. Their capacity as evaluation office has remained stagnant; they do cost recovery and they are quite agile.

GCF: the evaluation policy should be general, because all are not expert, also because they do not want the board to micro-manage evaluation policies. Learning and accountability are important. There is also a need for a very good communication policy inside Evaluation. The budget should defined at an early stage.

UNESCO: evaluation to improve performance of the evaluation, go with the grain, if a policy does not fit within an evaluation framework, it will not be used; learning, what about mechanisms to learn from management response.

Annex 1 – List of EPE participants

No.	Surname	Name	Organization	E-mail
1	Dillon	Neil	ALNAP	n.dillon@alnap.org
2	Goldman	Maria	COE	Maria.goldman@coe.int
3	Terbeck	Rica	COE	Rica.terbeck@coe.int
4	Van Hemelrijck	Adinda	Collaborative Impact	adinda@piala.org
5	Nega	Eskedar	ECA	nega.uneca@un.org
6	Lyth	David	EHM	davidwlyth@gmail.com
7	Herzog	Felix	ESCWA	herzogf@un.org
8	Zgambo	Olive	FAO	Olive.Zgambo@fao.org
9	Assaf	Jenin	FAO	Jenin.Assaf@fao.org
10	Awabdeh	Omar	FAO	Omar.Awabdeh@fao.org
11	Bayryyev	Serdar	FAO	Serdar.bayryyev@fao.org
12	Beasca	Joel	FAO	Joel.Beasca@fao.org
13	Belli	Luisa	FAO	luisa.belli@fao.org
14	Beutling	Anne	FAO	Anne.beutling@fao.org
15	Bonomi	Genny	FAO	genny.bonomi@fao.org
16	Bottamedi	Federica	FAO	Federica.Bottamedi@fao.org
17	Bruno	Marta	FAO	Marta.Bruno@fao.org
18	Celse-L'Hoste	Margareth	FAO	margareth.celselhoste@fao.org
19	Citerin	Dylan	FAO	Dylan.Citerin@fao.org
20	Coccia	Maria Federica	FAO	Federica.Coccia@fao.org
21	Cossee	Olivier	FAO	Olivier.Cossee@fao.org
22	Duah	Maame	FAO	Maame.duah@fao.org
23	Garcia	Harvey	FAO	harvey.garcia@fao.org
24	Holst	Sara	FAO	Sara.holst@fao.org
25	Igarashi	Masa	FAO	Masahiro.Igarashi@fao.org
26	Johnson	Vinitha	FAO	Vinitha.johnson@fao.org
27	Khalid	Arwa	FAO	Arwa.Khalid@fao.org
28	Khan	Mikal	FAO	Mikal.Khan@fao.org
29	Lappo	Alena	FAO	Alena.lappo@fao.org
30	Larmoyer	Aurelie	FAO	Aurelie.Larmoyer@fao.org
31	Lorvao	Clement	FAO	lorvao@outlook.com
32	Mirulla	Renata	FAO	Renata.mirulla@fao.org
33	Monforte	Lavinia	FAO	lavinia.monforte@fao.org
34	Monsour	Veridiana	FAO	Veridiana.mansourmendes@fao.org
35	Ould Abdallahi	Ahmedou	FAO	Ahmedou@hotmail.com
36	Solal-Celigny	Amelie	FAO	Amelie.SolalCeligny@fao.org
37	Tarazona	Carlos	FAO	Carlos.Tarazona@fao.org
38	Tessitore	Savina	FAO	savina.tessitore@gmail.com
39	Yabuki	Nanae	FAO	Nanae.yabuki@fao.org
40	Baguma	Emmanuella	GAVI	ebaguma@gavi.org
41	Moreland	Leslie	GAVI	lmoreland@gavi.org
42	Puri	Jyotsna	GCF	jpuri@gcfund.org
43	Carugi	Carlo	GEF	ccarugi@thegef.org
44	Uitto	Juha	GEF	juitto@thegef.org

No.	Surname	Name	Organization	E-mail
45	Kojoyan	Seda	Global Fund to Fight AIDS	seda.kojoyan@theglobalfund.org
46	Jaiyen	Kamolmas	IAEA	k.jaiyen@iaea.org
47	Thomas	Leslie	IAEA	l.thomas@iaea.org
48	Farice	Gugsa Yimer	ICAO	gfarice@icao.int
49	Abi Khalil	Diana	IFAD	d.abikhalil@ifad.org
50	Carbon	Michael	IFAD	m.carbon@ifad.org
51	Deshpande	Chitra	IFAD	c.deshpande@ifad.org
52	Federkeil	Johannes	IFAD	j.federkeil@ifad.org
53	Felloni	Fabrizio	IFAD	f.felloni@ifad.org
54	Garcia	Oscar	IFAD	o.garcia@ifad.org
55	Khaira	Hansdeep	IFAD	H.khaira@ifad.org
56	Kodjo	Max Kouessi	IFAD	k.kodjo@ifad.org
57	Lumaldo	Nicoletta	IFAD	n.lumaldo@ifad.org
58	Nakai	Fumiko	IFAD	f.nakai@ifad.org
59	Perch	Catrina	IFAD	c.perch@ifad.org
60	Piccolella	Antonella	IFAD	a.piccolella@ifad.org
61	Somma	Simona	IFAD	s.somma@ifad.org
62	Gillijns	Josse	IFRC	josse.gillijns@ifrc.org
63	Bavitch	Nathalie	ILO	bavitch@ilo.org
64	Vidal	Patricia	ILO	vidalhurtado@ilo.org
65	Sibanda	Adeline	IOCE	troparg@yahoo.com
66	Franzetti	Christophe	IOM	cfranzetti@iom.int
67	Paducel	Anca	IOM	apaducel@iom.int
68	Jimenez Pont	Michel	ITC	jimenez@intracen.org
69	Castells	Nuria	JIU	nuria.castells@un.org
70	Helck	Stefan	JIU	Stefan.helck@un.org
71	Jankauskas	Vytautas	LMU	vytautasjan@gmail.com
72	Soderini	Ludovica	MEF	Ludovica.soderini@mef.gov.it
73	Osembo	Sarah	OCHA	OSEMBO@UN.ORG
74	Saiz Omenaca	Victoria	OCHA	saiz-omenaca@un.org
75	Monroy	Sabas	OHCHR	smonroy@ohchr.org
76	Craft	Michael	OIOS	michael.craft@un.org
77	Guo	Eddie	OIOS	guoy@un.org
78	Muhadinovic	Milena	OIOS	muhadinovic@un.org
79	Rao	Srilata	OIOS	Rao3@un.org
80	Wieczerzak	Lukasz	OPCW	Lukasz.wieczerzak@opcw.org
81	Boschi	Melisso	Italian Senate	melisso.boschi@senato.it
82	Lambiase	Fortunato	Italian Senate	fortunato.lambiase@senato.it
83	Pegurri	Elisabetta	UNAIDS	pegurrie@unaids.org
84	Fyfe	Andrew	UNCDF	andrew.fyfe@uncdf.org
85	Legrand	Christophe	UNCDF	Christophe.legrand@uncdf.org
86	Bajwa	Madeeha	UNCTAD	madeeha.bajwa@unctad.org
87	Sofroni	Janna	UNCTAD	janna.sofroni@unctad.org
88	Acosta	Natalia	UNDP	Natalia.acosta@undp.org
89	Bless	Youri	UNDP	youri.bless@undp.org
90	Cole	Conception	UNDP	concepcion.cole@undp.org
91	Guinot	Mar	UNDP	mar.guinot@undp.org
92	Hauge	Arild	UNDP	arild.hauge@undp.org

No.	Surname	Name	Organization	E-mail
93	Jones	Richard	UNDP	Richard.jones@undp.org
94	Naidoo	Indran	UNDP	Indran.naidoo@undp.org
95	Naik	Shabbir	UNDP	Shabbir.naik@undp.org
96	Ouchi	Fumika	UNDP	Fumika.ouchi@undp.org
97	Wong	Catherine	UNDP	Catherine.wong@undp.org
98	Zhang	Jin	UNDP	Jin.zhang@undp.org
99	Rai	Ninamma	UNDPA	Ninamma.rai@un.org
100	Igartua	Daniel	UNDSS	igartua@un.org
101	Spilsbury	Michael	UNEP	michael.spilsbury@un.org
102	Wildish	Janet	UNEP	wildishjl@yahoo.com
103	Frueh	Susanne	UNESCO	s.frueh@unesco.org
104	Piric	Amir	UNESCO	a.piric@unesco.org
105	Adelakin	Olugbemiga	UNFPA	adelakin@unfpa.org
106	Alam	Mahbub	UNFPA	malam@unfpa.org
107	Chambel	Alexandra	UNFPA	chambel@unfpa.org
108	Daoudi	Hicham	UNFPA	daoudi@unfpa.org
109	Lenci	Sergio	UNFPA	slenci@unfpa.org
110	Mattson	Susanne	UNFPA	mattsson@unfpa.org
111	Mora-Brito	Daniel	UNFPA	mora-brito@unfpa.org
112	Segone	Marco	UNFPA	segone@unfpa.org
113	Barugahare	Martin	UN-Habitat	Martin.Barugahare@un.org
114	Bech	Susanne	UN-Habitat	Susanne.bech@un.org
115	Omondi	Lucy	UN-Habitat	Lucy.Omondi@un.org
116	Bonino	Francesca	UNHCR	bonino@unhcr.org
117	Bell	Lori	UNICEF	lbell@unicef.org
118	Laryea-Adjei	George	UNICEF	glaryeaadjei@unicef.org
119	Ocampo-Cobos	Ada	UNICEF	aocampo@unicef.org
120	Sabatini	Fabio	UNICEF	fsabatini@unicef.org
121	Dolun	Muge	UNIDO	u.dolun@unido.org
122	Guarnizo	Javier	UNIDO	J.Guarnizo@unido.org
123	Le	Thuy Thu	UNIDO	t.le@unido.org
124	Boyer	Brook	UNITAR	Brook.boyer@unitar.org
125	Koke	Katinka	UNITAR	Katinka.koke@unitar.org
126	Meiners	Olivia	UNITAR	Olivia.meiners@unitar.org
127	Asenjo Ruiz	Carlos Andres	UNODC	Carlos.asenjoruiz@un.org
128	Aston	Katherine	UNODC	katherine.aston@un.org
129	Gunnarsson	Charlotte	UNODC	charlotte.gunnarsson@un.org
130	Stryk	Robert	UNRWA	r.stryk@unrwa.org
131	Sniukaite	Inga	UNWOMEN	Inga.sniukaite@unwomen.org
132	Suarez	Maria Isabel	UNWOMEN	Isabel.suarez@unwomen.org
133	Tassew	Messay	UNWOMEN	Messay.tassew@unwomen.org
134	Benoit	Elise	WFP	elise.benoit@wfp.org
135	Borlini	Roberto	WFP	Roberto.borlini@wfp.org
136	Cook	Andrea	WFP	andrea.cook@wfp.org
137	Desole	Ramona	WFP	Ramona.desole@wfp.org
138	Duffy	Gabrielle	WFP	Gaby.duffy@wfp.org
139	Fernandez	Diego	WFP	diego.fernandez@wfp.org
140	Figus	Elena	WFP	Elena.figus@wfp.org

No.	Surname	Name	Organization	E-mail
141	Flentge	Jaqueline	WFP	Jacqueline.flentge@wfp.org
142	Garcia	Alberto	WFP	alberto.garciahernandez@wfp.org
143	Habtemariam	Dawit	WFP	dawit.habtemariam@wfp.org
144	Honjo	Mari	WFP	mari.honjo@wfp.org
145	Hurlen	Marte	WFP	marte.hurlen@wfp.org
146	Mcwhinney	Deborah	WFP	deborah.mcwhinney@wfp.org
147	Nguyen	Tram	WFP	tram.nguyen@wfp.org
148	Pappalepore	Giulia	WFP	giulia.pappalepore@wfp.org
149	Pini	Giancarlo	WFP	giancarlo.pini@wfp.org
150	Reynolds	Michael	WFP	mike@mreynolds.net
151	Spacca	Arianna	WFP	arianna.spacca@wfp.org
152	Thoulouzan	Julie	WFP	Julie.thoulouzan@wfp.org
153	Zelada	Federica	WFP	federicazelada@gmail.com
154	Bettighofer	Simon	WHO	Bettighofers@who.int
155	Drayton	Carol	WHO	draytonc@who.int
156	Larizgoitia	Itziar	WHO	larizgoitiai@who.int
157	Luzot	Anne-Claire	WHO	luzota@who.int
158	Renganathan	Elil	WHO	renganathane@who.int
159	Sivasankara Kurup	Anand	WHO	sivasankarakurupa@who.int
160	Caira	Celine	WIPO	celinevcaira@gmail.com
161	Engelhardt	Julia	WIPO	Julia.engelhardt@wipo.int
162	Ruiz	Adan	WIPO	adan.ruizvillalba@wipo.int

Annex 2 – EPE Agenda

Tuesday 8 May			
09.00-10.15	**Opening session*** - Sheikh Zayed Center		
10.15-11.00	**Group photo** and **Coffee break** - Atrium		
11.00-12.30	Session 1.1 • **Quality Assurance beyond the checklist: how can we ensure the reports are evidence-based?** **OSCAR GARCIA**, IFAD **Philippines Room** C277	Session 2.1 • **We ask others to demonstrate impact: how can we demonstrate the impact of our evaluations?** **DEBORAH MCWHINNEY**, WFP **Ethiopia Room** C285	Session 3.2 • **Making the best use of Theories of Change in Evaluations** **MICHAEL CARBON**, IFAD **Iraq Room** A235
12.30-14.00	**Lunch**		
13.15-14.00	Lunch-time seminar • **Evaluating Capacity Development: sharing experiences** **LAVINIA MONFORTE**, FAO **Iraq Room 2** A235	**Professionalization of Evaluation Working Group Meeting – the way forward** **JACQUELINE FLENTGE** (WFP) and **SUSANNE MATTSSON** (UNFPA) **Nigeria Room** C215	
14.00-15.30	Session 1.4 • **How can we ensure a gender focused evaluation?** **INGA SNIUKAITE**, UNWOMEN **Philippines Room** C277	Session 2.3 • **Cost of evaluations: how to rationalize the use of resources for evaluations?** **AMÉLIE SOLAL-CÉLIGNY**, FAO **Ethiopia Room** C285	Session 3.4 • **ICT for data collection and analysis in evaluations** **KATINKA KOKE**, UNITAR **Iraq Room** A235
15.30-16.00	**Coffee break: video projected** (Myanmar FAO Country Programme Evaluation, 20') - **Iraq Room** A235		
16.00-17.30	Session 1.3 • **As evaluation managers, we all face similar challenges: let's share them and learn from them** **JULIE THOULOUZAN**, WFP **Iraq Room** A235	Session 2.4 • **Evaluation and hindsight: assessing past interventions against yesterday's standards or today's wisdom and criteria** **FABRIZIO FELLONI**, IFAD **Ethiopia Room** C285	Session 3.1 • **No baseline: how to measure impact?** **ANNA HENTTINEN**, WFP **Philippines Room** C277
19.00	**Self-paid dinner** - Osteria Circo - Via dei Cerchi, 79		

Wednesday 9 May			
09.00-10.30	Session 1.2 • **Following up on recommendations: what are the successful practices?** **JULIA ENGELHARDT**, WIPO **Iraq Room** A235	Session 2.6 • **Full Disclosure? What evaluators don't write about** **OLIVIER COSSÉE**, FAO **Ethiopia Room** C285	Session 3.6 • **Emerging principles and lessons for the development of organizational evaluation policies** **JYOTSNA PURI (JO)**, GCF **Philippines Room** C277
10.30-11.00	**Coffee break: video projected** (Niger FAO Country Programme Evaluation, 10') - **Iraq Room** A235		
11.00-12.30	Session 1.5 • **The role of donors in evaluations: how should we involve them and how to manage expectations** **AMIR PIRIC**, UNESCO **Ethiopia Room** C285	Session 2.2 • **In a "post-truth" era, how can we evaluators ensure we adapt and remain heard?** **ADAN RUIZ VILLALBA**, WIPO **Philippines Room** C277	Session 3.5 • **Evaluating policy support** **VERIDIANA MANSOUR**, FAO **Iraq Room** A235
12.30-14.00	**Lunch**		
13.15-14.00	Lunch-time seminar • **Decentralized Evaluation and National Evaluation Capacity: hands-on experiences from the field** **ADAN RUIZ VILLALBA**, WIPO and **ALEXANDRA CHAMBEL**, UNFPA **Iraq Room 2** A235	**Ethics and Code of Conduct Guidance Task Force meeting – the way forward** **GABY DUFFY**, WFP **Nigeria Room** C215	
14.00-15.30	Session 1.6 • **Procuring UN evaluators externally: learning from current practices and suggestions for improvement** **ANDREW FYFE**, UNCDF **Iraq Room** A235	Session 2.5 • **What does independence mean today – A debate!** **JYOTSNA PURI (JO)**, GCF **Ethiopia Room** C285	Session 3.3 • **Country Programme Evaluations - methodological challenges** **DAWIT HABTEMARIAM**, WFP **Philippines Room** C277
15.30-16.00	**Coffee break: presentation of the "Unite Evaluation" tool** by **KATHARINA KAYSER**, UNODC - **Iraq Room** A235		
16.00-17.30	**Wrap-up session*** - **Iraq Room** A235		

* these sessions will be webstreamed

Annex 3 – Results of post-EPE survey

PDS trainings

This innovation was massively supported by the participants according to the survey with only three participants declaring they would not like the experience to be repeated. Most of the participants enjoyed having a **complement to the interactive sessions**. Indeed, by nature, the EPE sessions necessarily have a more loose structure and, as they are shorter, cannot **deepen topics** as much as the PDS. This **mix of formal and more informal learning** was apparently duly appreciated. Comments from diverse participants show that while some participants were introduced to a new topic, some other could refresh their knowledge, which contributed to creating a common basis among evaluation offices.

Besides, as it was the first time that such a process was incorporated to the EPE, useful suggestions of improvement were made. The time constraint seems to be one of the major point of discord since an important number of comments emphasized on the fact that the trainings could have lasted longer to deepen the topics. On the other hand, few participants are reticent to adding one day to the EPE as it might become problematic for some institutions.

Regarding the content, it was argued that the training could have been **less theoretical and more practical**. Indeed, the question of **applicability** appears as a major concern for participants among whom many called for trainings that adapt to the evaluation context. Also concerning the content, some participants deplored that the trainings were mostly tailored for managers and staff, and would have enjoyed trainings for different backgrounds, such as junior participants. Overall, the selection of topics appears as a recurrent matter of concern in comments, even though participants were solicited before the EPE through needs assessments.

Eventually, a participants suggested to engage trainers from recognized institutions instead of consultants to facilitate the follow-up contacts between trainees and alumni.

Logistics

Conveners and collaborators were overall satisfied with the support provided before and during the EPE. They did not leave any comments for improvements on that topic.

Participants were also satisfied with the preparation of the EPE. The logistical note was consulted by more than 75% of the participants, but some topics were considered less relevant, namely the "online community of practice part" and the "documentation (link to UNEG website)". Besides, participants would have enjoyed to receive the agenda earlier.

The rooms selected were considered correct by participants. The absence of windows and issues of acoustic in the Philippines and Ethiopia, and the small size of the Nigeria room, were pointed out as drawbacks.

Some comments argued for some changes regarding the conveners and collaborators selection. It was notably stated that the EPE should involve more staff from regional and country levels. Moreover, it was suggested to reconsider the relation between conveners and collaborators which was not clear enough according to some participants. The possibility of having **co-conveners** was suggested twice, as it would also allow to compensate the potential absence of a convener.

Interestingly, lead conveners (86%) supported the principle of making a video to introduce their session more than the participants (72%). Comments noted that videos added little value in terms of content and few participants would say that they chose their sessions based on the videos. However, it permitted participants to associate names with faces which made contacts easier. Some expressed concerns on the fact that the video might be confounding as some conveners could have more skills than others in front of the camera, and that some sessions could therefore be devaluated for issues of communication.

Concerning the organisation of the sessions. Participants were generally satisfied with the number and duration of the sessions, and with the coffee breaks. Opinions were somewhat more mitigated on the duration and relevance of the lunch time seminars, for which some participants expressed their difficulties to attend for time constraints. The social diner was considered as a relevant opportunity.

Some participants commented however that the **time allocated to the sessions** did not allow to reach the depth in the discussion and would advise to reduce their number. Another solution would be to have 1.5/2 hours sessions for some topics, and 4 hours sessions for others that warrant more in-depth learning.

EPE sessions and AGM

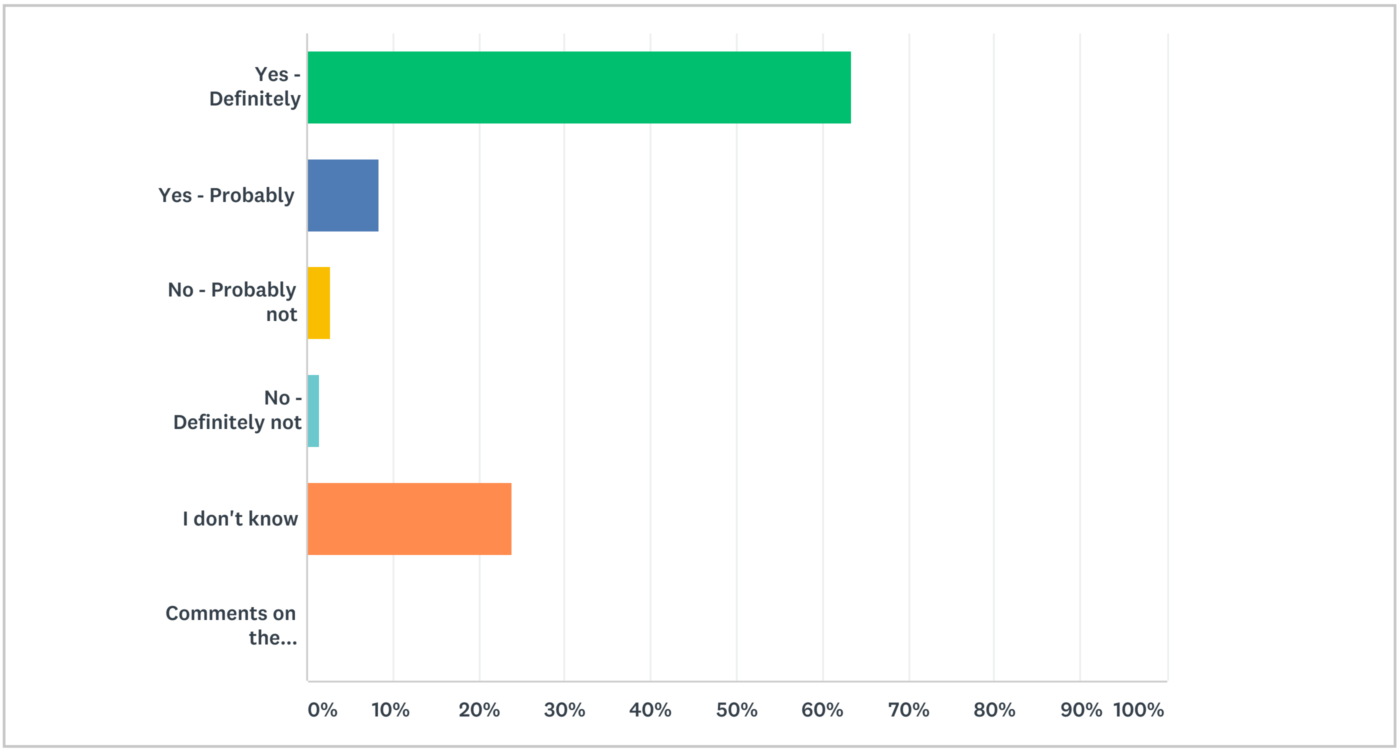
Sessions with interactive modalities were often appreciated by participants who noted that the **no-power point rule should be kept** (for 90% of them). **World cafés, Oxford-style debates, and role plays** received important support from the participants. The move from traditional panel presentations to more interactive and engaging modalities was also saluted in comments. As expressed earlier regarding the PDS, a category of participants, difficult to measure as it is based on few comments, consider that there is little value added in constantly sharing experiences over similar subjects, and would appreciate **more focus on masterclasses**. The **time allocated to masterclass and exchange sessions** is probably what concerned participants the most in the survey.

Topics that interested participants the most were related to the following up of recommendations, the data analysis, the independence and the cost of evaluations. Interesting comment proposed to **engage outsiders, specialist of the dedicated subjects**, for some sessions in order to raise even more the quality of the debates.

Concerning the AGM, the length (two days) was considered appropriate by the participants. The way it was organized also allowed to review progress and decide on strategies and work areas for the following years according to the participants. Some comments claimed for a **more informal** modality, somewhat similar to the EPE sessions. Regarding the MTR, participants stressed its utility but some think that it would have been interesting to hire the two consultants who conducted it to foster discussions.

Q1 Was the Professional Development Seminar an interesting opportunity and would you like it to be repeated for the next UNEG week?

Answered: 71   Skipped: 2



ANSWER CHOICES	RESPONSES	
Yes - Definitely	63.38%	45
Yes - Probably	8.45%	6
No - Probably not	2.82%	2
No - Definitely not	1.41%	1
I don't know	23.94%	17
Comments on the Professional Development Seminar concept	0.00%	0
TOTAL		71

#	COMMENTS ON THE PROFESSIONAL DEVELOPMENT SEMINAR CONCEPT	DATE	
	There are no responses.		

Q2 Comments on the PDS concept

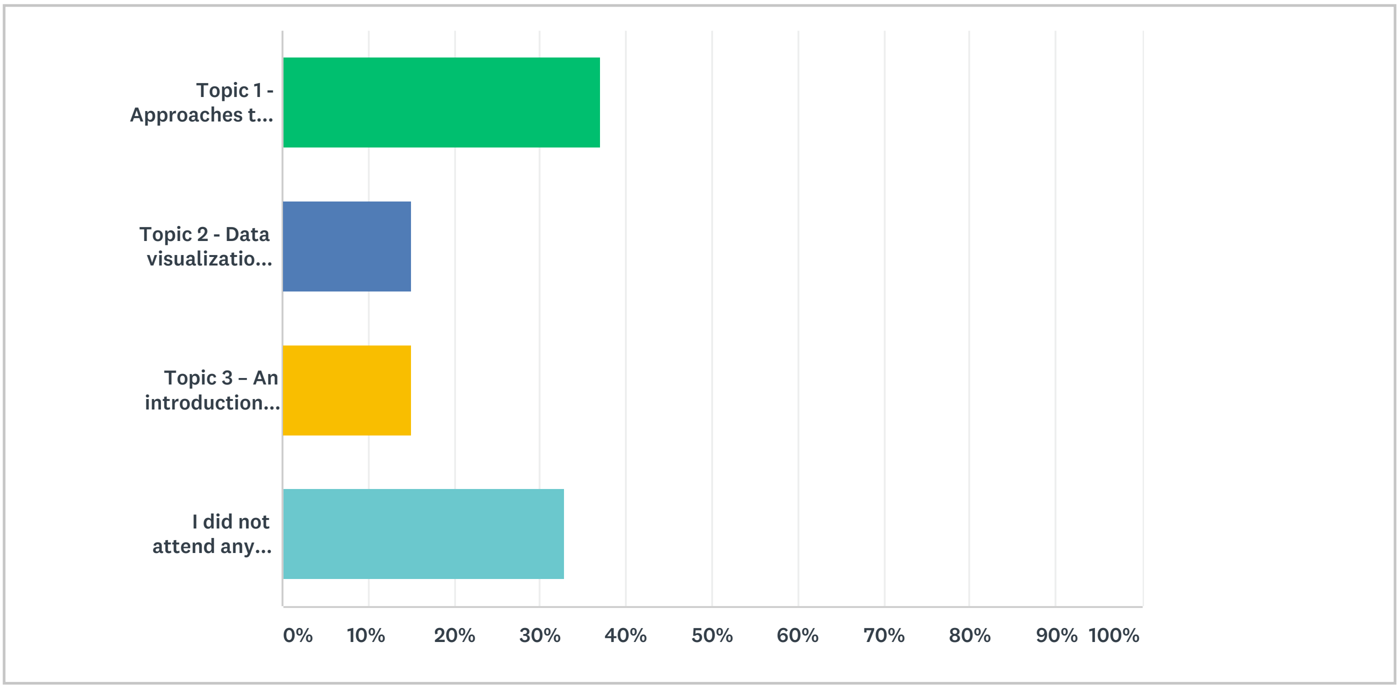
Answered: 38 Skipped: 35

#	RESPONSES	DATE
1	I did not attend.	6/6/2018 8:36 PM
2	Excelent value added topics and execution	6/5/2018 3:01 PM
3	I thought it was exceptionally useful and I feel like I benefited from the learning environment of the PDS more than I did from the EPE sessions which can vary greatly in their value.	6/5/2018 11:14 AM
4	It was useful and well prepared	6/5/2018 11:12 AM
5	Worked very well to have 3 in-depth topics to choose from (although they were each interesting)!	6/5/2018 10:01 AM
6	It was the most interesting section of the entire EPE. I encourage to offer more of this	6/5/2018 9:59 AM
7	it was very inspiring and innovative!!!	6/5/2018 9:50 AM
8	If it is possible, would be nice to have a possibility to attend more than one PDS and not to choose.	6/5/2018 9:16 AM
9	n.a.	6/5/2018 8:35 AM
10	Great opportunity to refresh or become introduced to a concept. One day is not sufficient as training however so important to explain how more can be done.	6/5/2018 8:32 AM
11	This is exactly what we needed. We need to update our knowledge in terms of new methods and practices and the PDS was very useful in this regards.	6/5/2018 8:13 AM
12	Idea is great but training needs to be more intense since it's such a short period...	6/4/2018 10:46 PM
13	innovative organization and full participation	6/4/2018 6:10 PM
14	Designed only for managers and not for more junior (doers staff)	6/4/2018 5:50 PM
15	The only problem is we cannot attend more than one PDS	6/4/2018 5:37 PM
16	An interesting concept to discuss important technical/ethical issues of evaluation.	5/30/2018 7:42 PM
17	The impact evaluation session was very good but the time allocated was inadequate to cover it	5/28/2018 3:30 PM
18	Could be extended to a 2 days workshop -especially if the topic is impact evaluation.	5/28/2018 2:41 PM
19	I did not attend the PDS due to a concurrent side meeting. I would like to share the feedback that I thought the concept was very good. Particularly appreciated the consultation ahead of selecting the themes.	5/28/2018 1:32 PM
20	Having 3 choices made it easy to find a relevant and interesting topic. It added a good mix of more formal learning with the informal EPE	5/25/2018 9:59 AM
21	the workshop format allowed for much more learnings than the conference format. the full day session including interactive and participative exercises were needed for reinforcing existing learning and acquiring new knowledge and skills. Great job!	5/25/2018 9:46 AM
22	The PDS was very interesting and useful . Efforts need to be made to make it less theoretical and more practical The one Inattended on impact evaluation was too theoretical and did not really adress practical challenges	5/25/2018 6:42 AM
23	I did not attend this session	5/24/2018 10:11 PM
24	I didn't participate.	5/24/2018 9:47 PM
25	I did not attend.	5/24/2018 1:51 PM
26	Andy Kirk was excellent in explaining the theory and practice of data visualization. It was extremely useful!	5/24/2018 1:18 PM
27	I liked the possibility of taking courses both on rather traditional methodologies and innovative	5/24/2018 8:16 AM

	ones.	
28	good idea. should be continued	5/24/2018 8:08 AM
29	Concept is good. Next time better to engage trainers from recognized institutions instead of individual consultants since this could facilitate follow-up contacts between trainees and alumni	5/23/2018 9:41 PM
30	I did not attend because the professionalization group I am part of, organized a round table at WFP the same day,	5/23/2018 6:47 PM
31	Popular met needs	5/23/2018 6:20 PM
32	I attended the impact evaluation PDS. it was excellent. pitty it was not possible to attend more than one pds.	5/23/2018 4:43 PM
33	I found the course on sense maker useful, but it did add a day to the UNEG which one may not be able to afford in all circumstances.	5/23/2018 4:42 PM
34	None	5/23/2018 4:13 PM
35	Great idea, but the training itself could have been more practical for evaluators to implement and less of a review of theory	5/23/2018 4:12 PM
36	Excellent idea to expand the focus of the Evaluation Week on UNEG members and their professional development and not only during the EPE. Also appreciated that the quality of the trading was so high (impact evaluation session). This makes it easier to justify being out of the office for the full week.	5/23/2018 4:12 PM
37	very good	5/23/2018 4:10 PM
38	good idea	5/23/2018 2:56 PM

Q3 Which training did you attend?

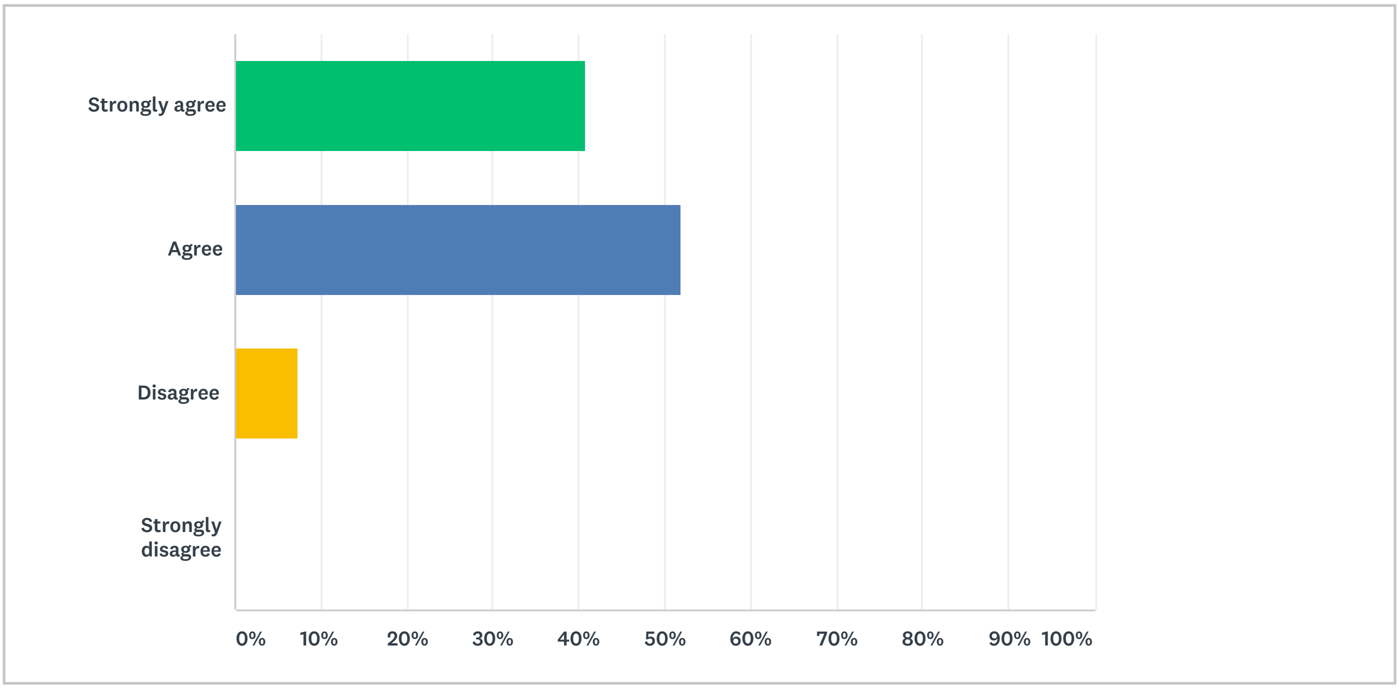
Answered: 73 Skipped: 0



ANSWER CHOICES	RESPONSES	
Topic 1 - Approaches to Impact Evaluation	36.99%	27
Topic 2 - Data visualization tools	15.07%	11
Topic 3 – An introduction to the practice of SenseMaker	15.07%	11
I did not attend any training	32.88%	24
TOTAL		73

Q4 The training "Approaches to Impact Evaluation" was useful for me and my organisation

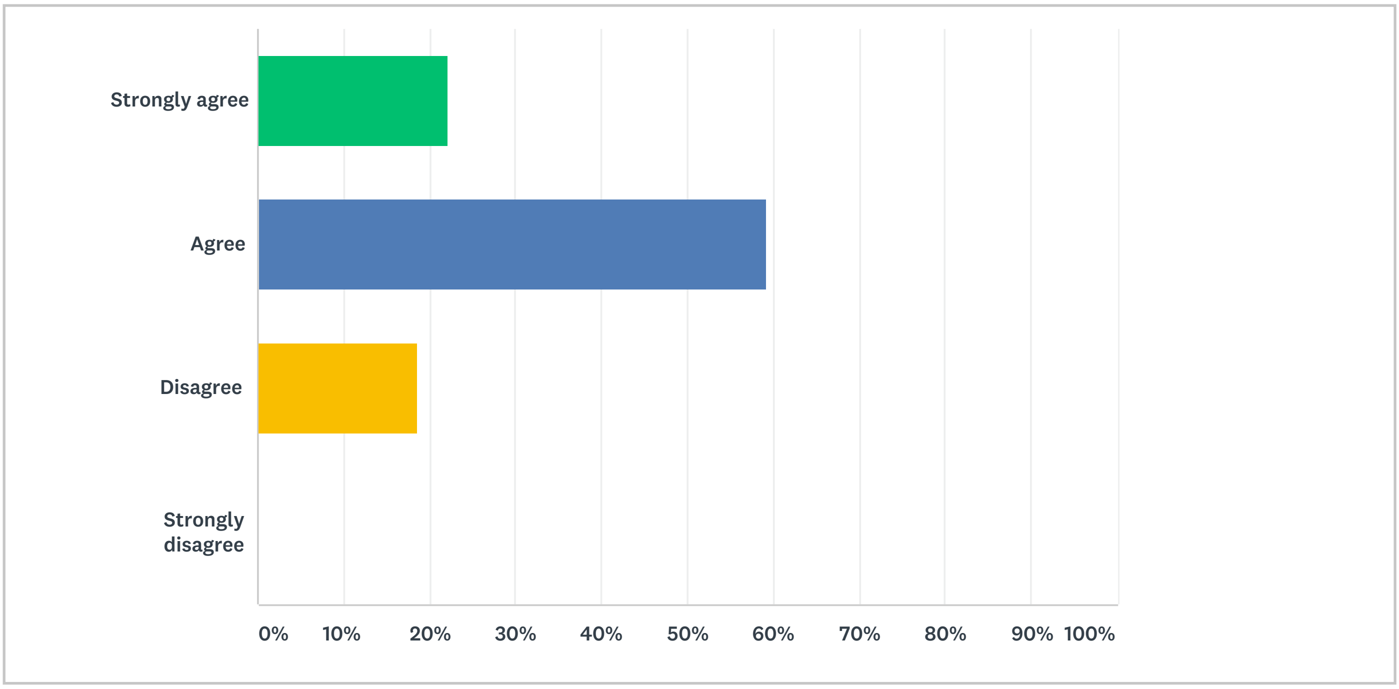
Answered: 27 Skipped: 46



ANSWER CHOICES	RESPONSES	
Strongly agree	40.74%	11
Agree	51.85%	14
Disagree	7.41%	2
Strongly disagree	0.00%	0
TOTAL		27

Q5 The training "Approaches to Impact Evaluation" has a high applicability for my work

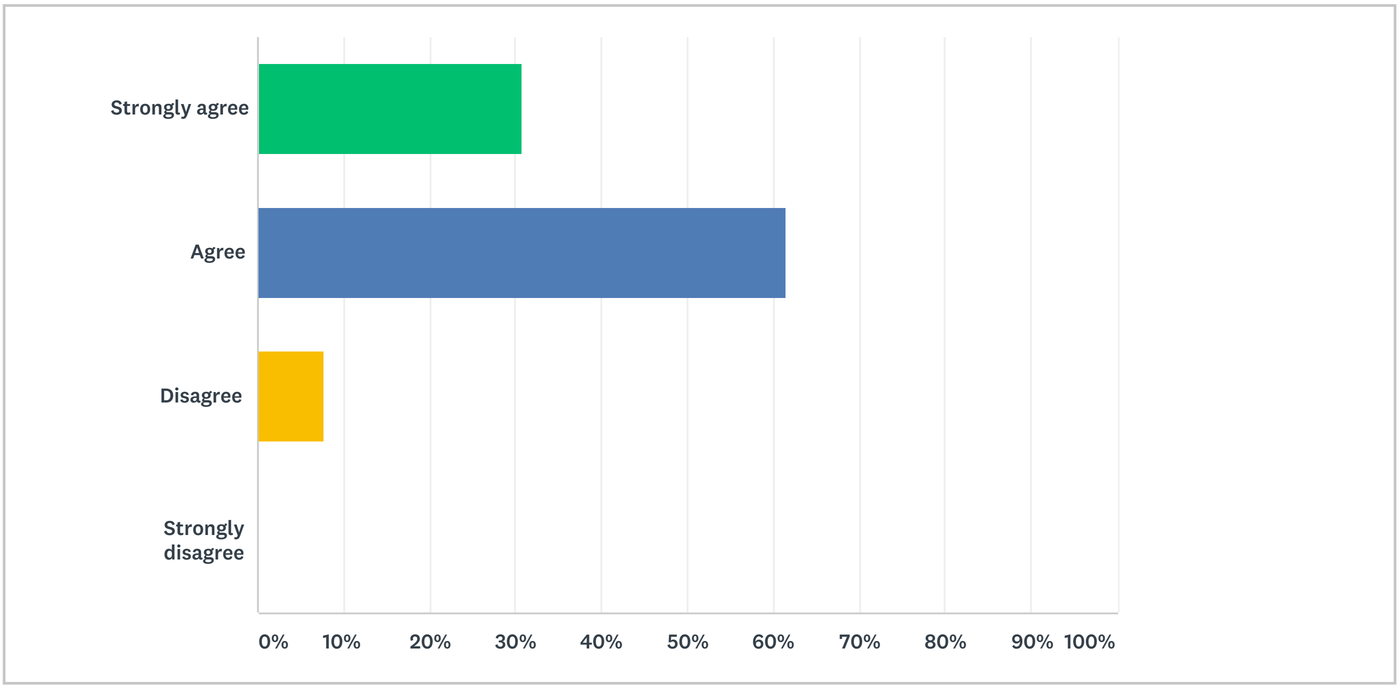
Answered: 27 Skipped: 46



ANSWER CHOICES	RESPONSES	
Strongly agree	22.22%	6
Agree	59.26%	16
Disagree	18.52%	5
Strongly disagree	0.00%	0
TOTAL		27

Q6 The modality of the training “Approaches to Impact Evaluation” was relevant

Answered: 26 Skipped: 47



ANSWER CHOICES	RESPONSES	
Strongly agree	30.77%	8
Agree	61.54%	16
Disagree	7.69%	2
Strongly disagree	0.00%	0
TOTAL		26

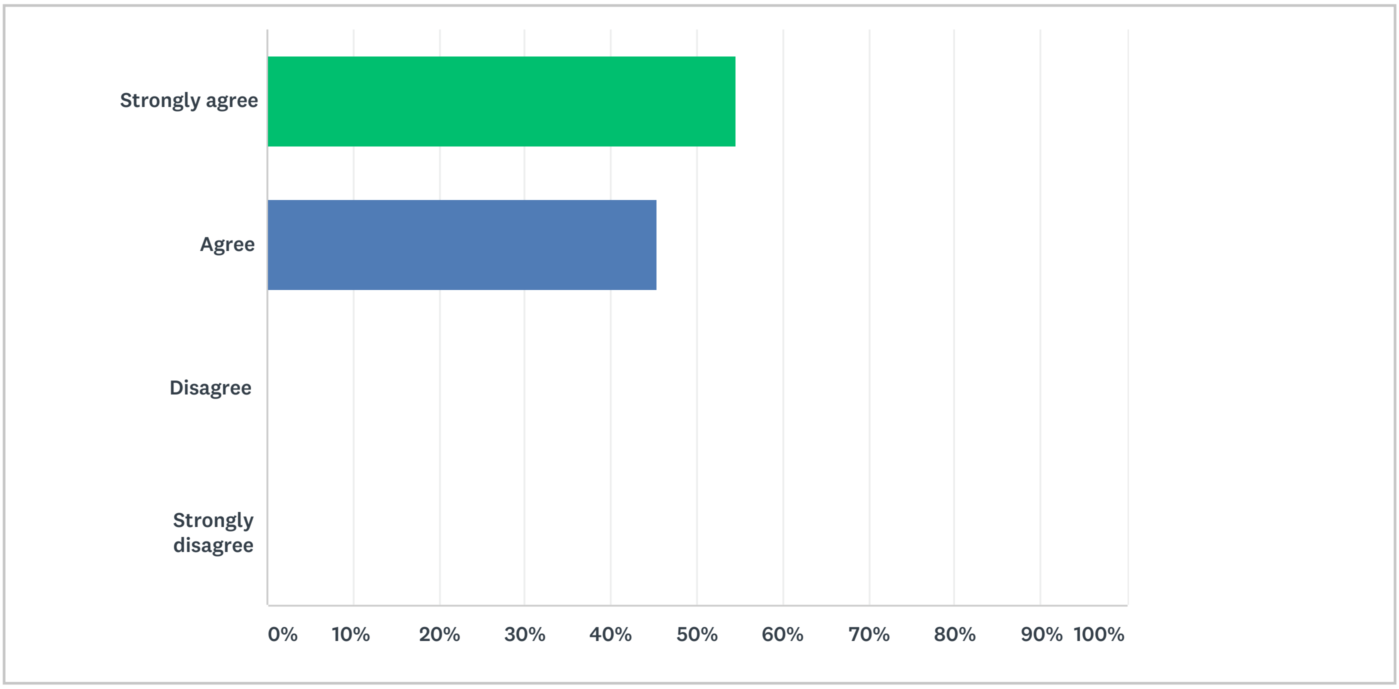
Q7 Comments or suggestions about the training “Approaches to Impact Evaluation”

Answered: 15 Skipped: 58

#	RESPONSES	DATE
1	I heard great feedback on each of the PDS events and I think it would be valuable to record them and share them with attendees. I imagine that speakers may not be too keen on that idea but I would have enjoyed benefitting from the other presentations as well. This class specifically was well presented given that it had three subject matter experts involved. There was a nice blend of presentation and practical exercises.	6/5/2018 11:17 AM
2	Well done. Nice combination of theory and praxis.	6/5/2018 10:01 AM
3	It was interesting. I suggest to seek more “masterclass” formats with technical specialists on the subject matters	6/5/2018 10:00 AM
4	One challenge was that some agencies just procure impact evaluation while others do it. I found the doing more interesting than the procuring. Perhaps for procuring a simple packet of info on can be provided with sample ToRs, costs, etc.	6/5/2018 8:35 AM
5	As mentioned, the target audience was the hiring manager rather than consultants or junior staff	6/4/2018 5:51 PM
6	Supported by practical experience. But specifically focused on projects but not impact of policy level works	5/28/2018 3:31 PM
7	More examples and “clinics” with experts in impact evaluation to discuss real “live” impact evaluations.	5/28/2018 2:42 PM
8	the exercise allowing to apply all theoretical aspects was excellent. more individual feedback would have been appreciated, but I understand time was tight. Overall, an excellent workshop, well structured and interesting!	5/25/2018 9:48 AM
9	the training was excellent, but the applicability in the evaluation of UN programmes in middle income countries is limited. we should explore and discuss more on how to apply rigorous impact assessment methodologies to, for instance, policy advocay initiatives.	5/23/2018 4:45 PM
10	It was not very practical for the scope of my organization, especially since RCT are not feasible	5/23/2018 4:14 PM
11	None	5/23/2018 4:13 PM
12	examples used are a bit over simplistic	5/23/2018 4:11 PM
13	it lacked interaction, was very generic and not adapted to our UN evaluation work. I would have preferred more preparation in understanding what kind of evaluations we do in the various UN agencies.	5/23/2018 2:58 PM
14	I appreciated the mix of theory and praxis.	5/23/2018 6:14 AM
15	More examples on how to apply methods to real evaluations	5/22/2018 5:18 PM

Q8 The training "Data visualisation tools" was useful for me and my organisation

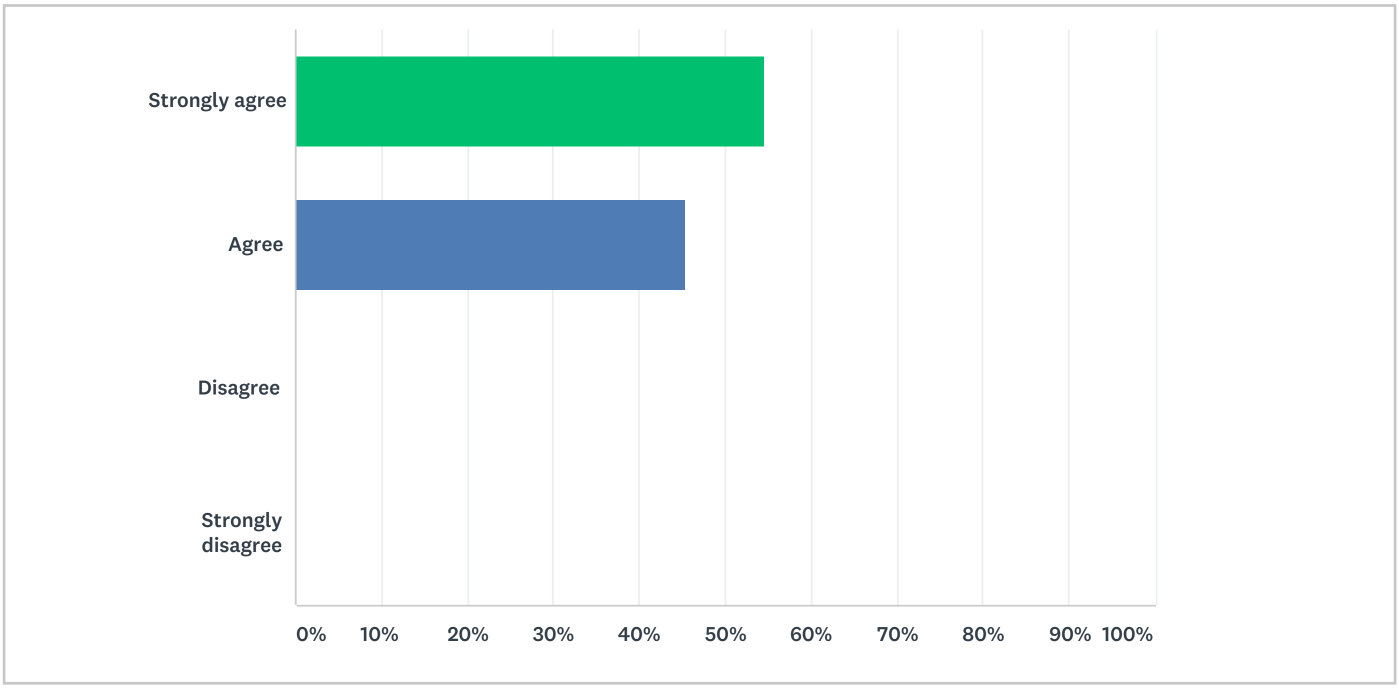
Answered: 11 Skipped: 62



ANSWER CHOICES	RESPONSES	
Strongly agree	54.55%	6
Agree	45.45%	5
Disagree	0.00%	0
Strongly disagree	0.00%	0
TOTAL		11

Q9 The training “Data visualisation tools” has a high applicability for my work

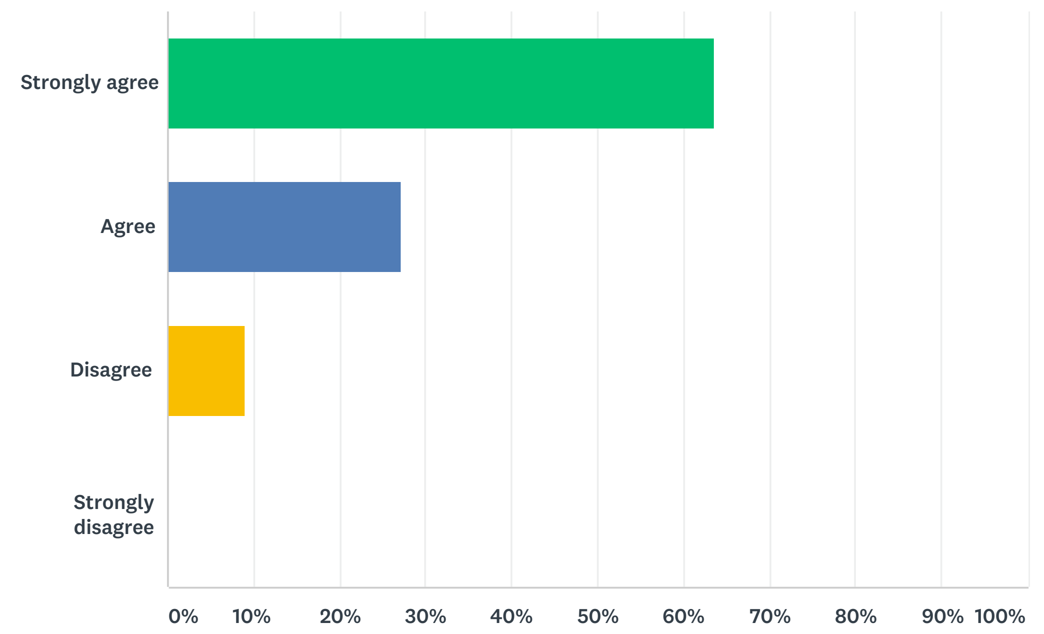
Answered: 11 Skipped: 62



ANSWER CHOICES	RESPONSES	
Strongly agree	54.55%	6
Agree	45.45%	5
Disagree	0.00%	0
Strongly disagree	0.00%	0
TOTAL		11

Q10 The modality of the training "Data visualisation tools" was relevant

Answered: 11 Skipped: 62



ANSWER CHOICES	RESPONSES	
Strongly agree	63.64%	7
Agree	27.27%	3
Disagree	9.09%	1
Strongly disagree	0.00%	0
TOTAL		11

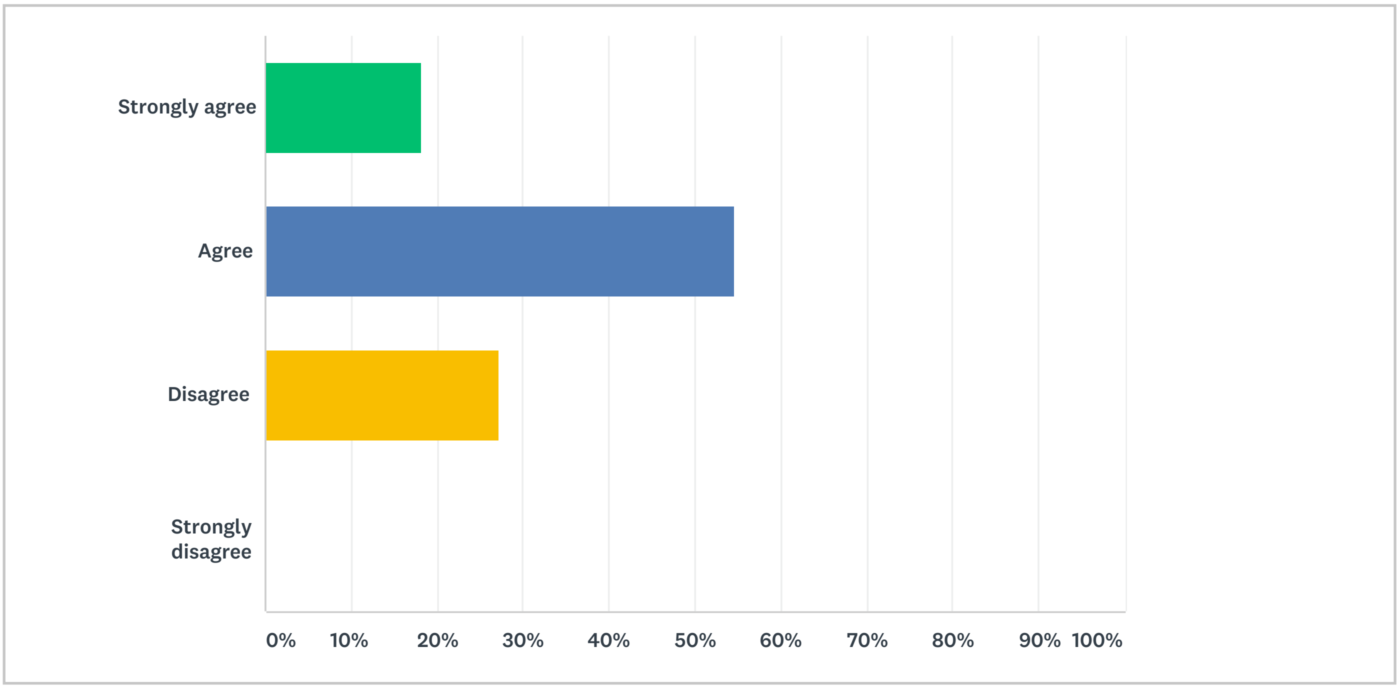
Q11 Comments or suggestions about the training “Data visualization tools”

Answered: 6 Skipped: 67

#	RESPONSES	DATE
1	More has to be done in this area but narrowing it down to the use od visualization in the specific evaluation field	6/5/2018 3:02 PM
2	It was very good. The time was too short but better one day then none.	6/5/2018 8:14 AM
3	Too much concentration on the looks and feel.of data visualization. Would have preferred training on how to better understand what to present and how.	6/4/2018 10:48 PM
4	topic rather a bit specialized, but interesting and good trainer	5/28/2018 11:50 AM
5	Very good facilitator/expert. It would be great for future training to link the subject to common practices in evaluation	5/23/2018 4:31 PM
6	Very good trainer, engaging session.	5/23/2018 11:49 AM

Q12 The training "An introduction to the practice of SenseMaker" was useful for me and my organisation

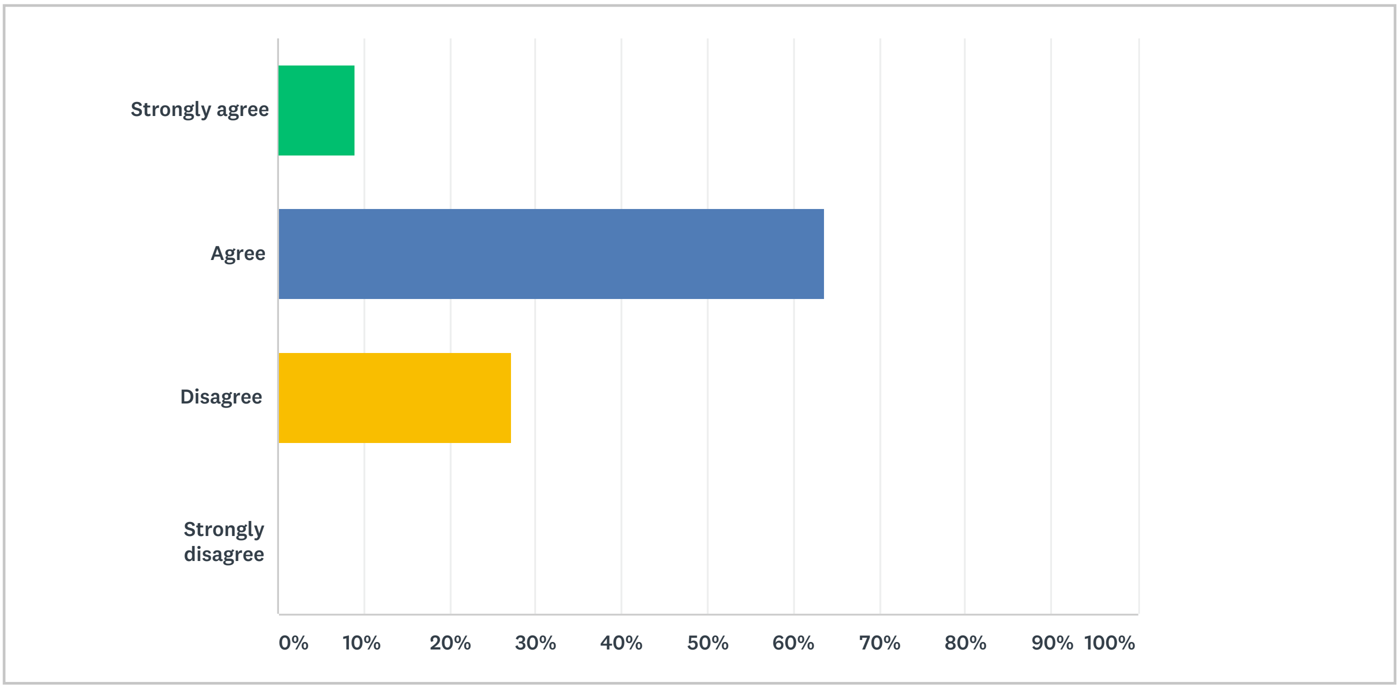
Answered: 11 Skipped: 62



ANSWER CHOICES	RESPONSES	
Strongly agree	18.18%	2
Agree	54.55%	6
Disagree	27.27%	3
Strongly disagree	0.00%	0
TOTAL		11

Q13 The training “An introduction to the practice of SenseMaker” has a high applicability for my work

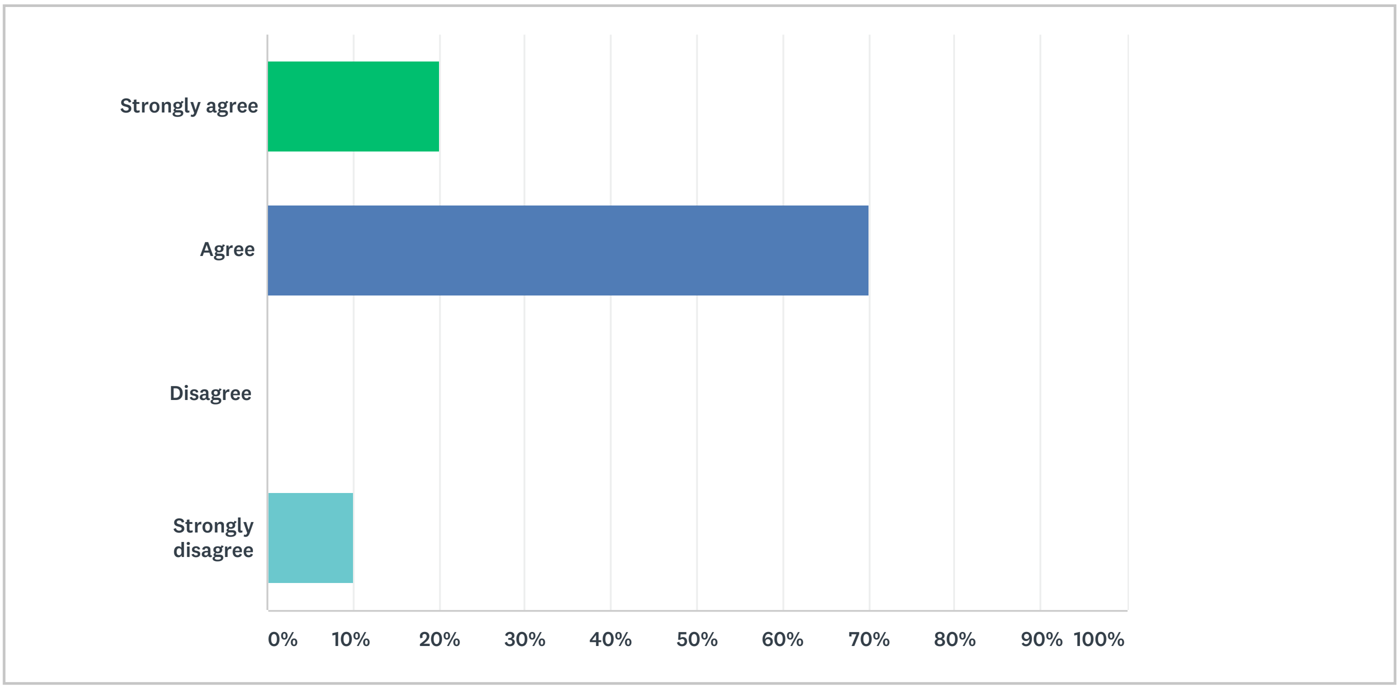
Answered: 11 Skipped: 62



ANSWER CHOICES	RESPONSES	
Strongly agree	9.09%	1
Agree	63.64%	7
Disagree	27.27%	3
Strongly disagree	0.00%	0
TOTAL		11

Q14 The modality of the training "An introduction to the practice of SenseMaker" was relevant

Answered: 10 Skipped: 63



ANSWER CHOICES	RESPONSES	
Strongly agree	20.00%	2
Agree	70.00%	7
Disagree	0.00%	0
Strongly disagree	10.00%	1
TOTAL		10

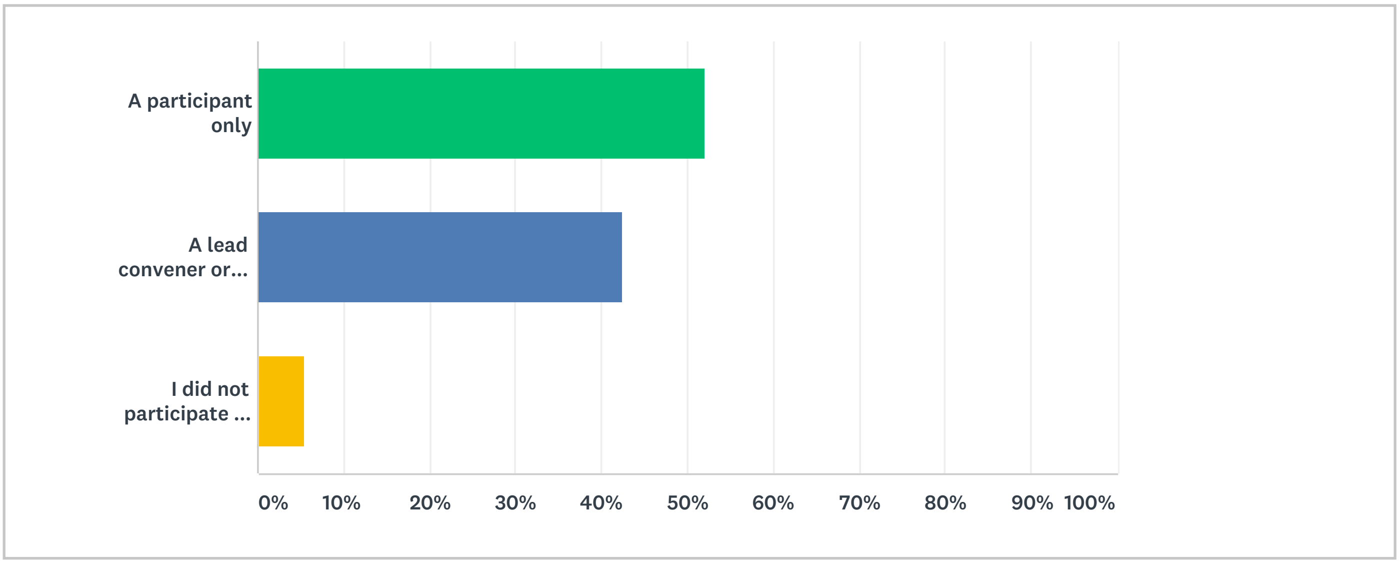
Q15 Comments or suggestions about the training “An introduction to the practice of SenseMaker”

Answered: 5 Skipped: 68

#	RESPONSES	DATE
1	Trainer was excellent and methodology very interesting.	6/5/2018 11:12 AM
2	Interesting approach but only applicable to big numbers and with use of software	6/5/2018 9:16 AM
3	It was just perfect, wouldn’t have done it differently.	5/24/2018 8:17 AM
4	Good overview; difficult to implement though	5/23/2018 9:42 PM
5	I found the afternoon session when the IFAD colleague presented his experience using the Sensemaker as the most interesting, I will advise that for future UNEG training workshops, the trainers should really adapt it to evaluation context. The morning session was extremely generic and not tailored to evaluation applicability	5/23/2018 4:13 PM

Q16 Did you participate in the EPE as:

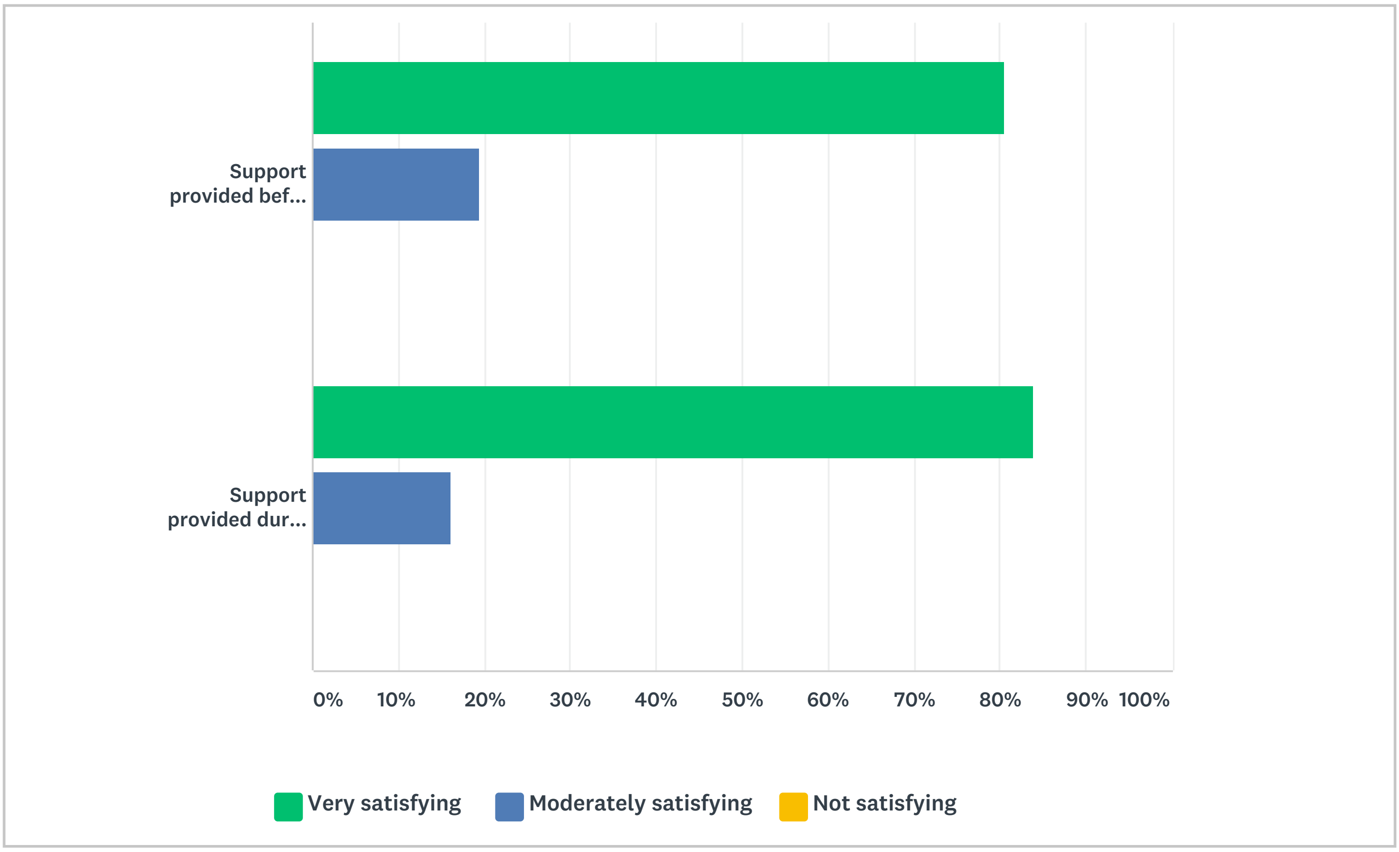
Answered: 73 Skipped: 0



ANSWER CHOICES	RESPONSES	
A participant only	52.05%	38
A lead convener or collaborator	42.47%	31
I did not participate in the EPE	5.48%	4
TOTAL		73

Q17 Please comment about the support provided by the EPE Committee to you as a lead convener or collaborator

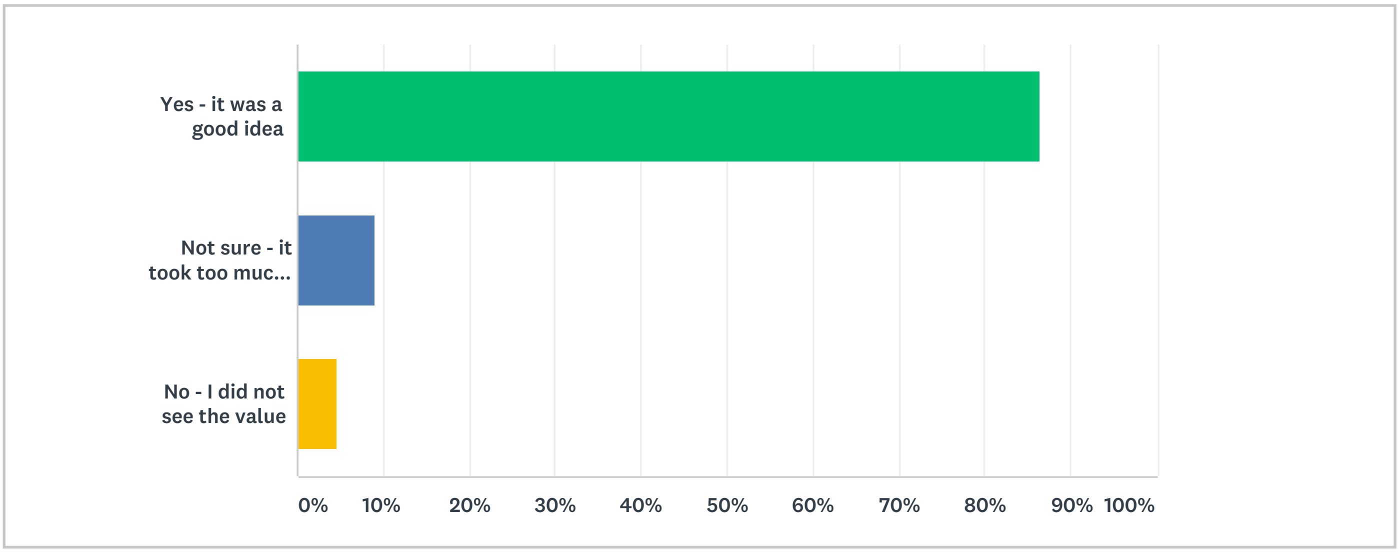
Answered: 31 Skipped: 42



	VERY SATISFYING	MODERATELY SATISFYING	NOT SATISFYING	TOTAL
Support provided before the EPE (preparation, planning, clarity of roles and responsibilities)	80.65% 25	19.35% 6	0.00% 0	31
Support provided during the EPE (logistics, agenda, clarity of expectations)	83.87% 26	16.13% 5	0.00% 0	31

Q18 As a lead convener, would you do the video again?

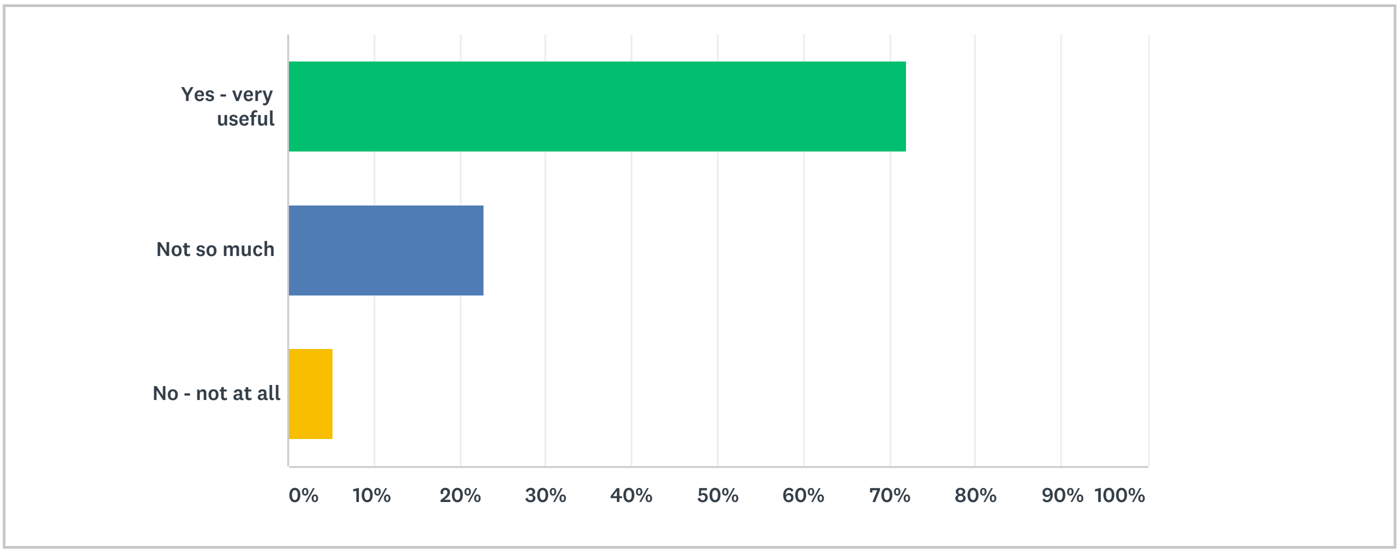
Answered: 22   Skipped: 51



ANSWER CHOICES	RESPONSES	
Yes - it was a good idea	86.36%	19
Not sure - it took too much time	9.09%	2
No - I did not see the value	4.55%	1
TOTAL		22

Q19 As a participant, did you find the EPE video useful?

Answered: 57 Skipped: 16



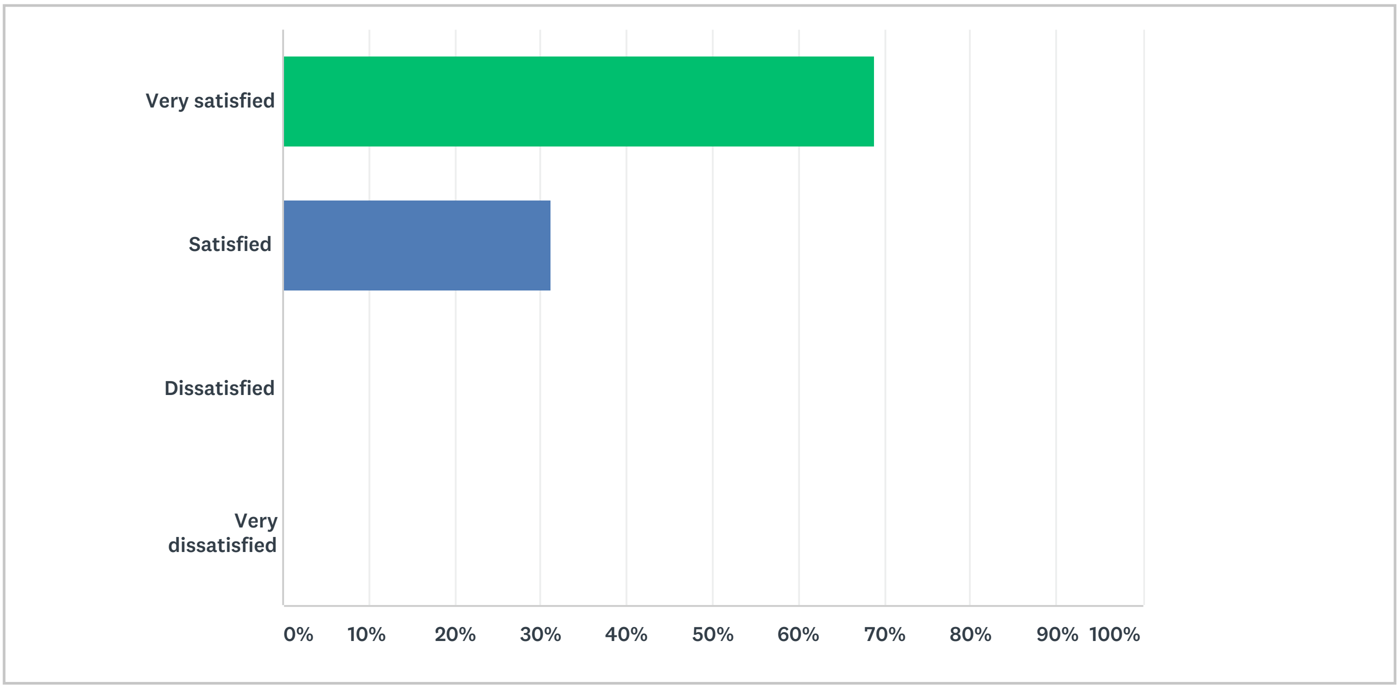
ANSWER CHOICES	RESPONSES	
Yes - very useful	71.93%	41
Not so much	22.81%	13
No - not at all	5.26%	3
TOTAL		57

#	COMMENTS ON THE EPE VIDEO	DATE
1	I thought it was fun to watch and it was of value but not sure I would say it was 'very useful.' It added value though. Of course producing a video like that may not always be possible for a host so I'd suggest a broader explanation in the printed handout - which was also quite helpful and a wonderful document.	6/5/2018 11:22 AM
2	Fine, but not needed	6/5/2018 11:13 AM
3	It was nicely filmed, but the usefulness was very much limited.It did not help to understand the content and focus of the sessions	6/5/2018 10:02 AM
4	What video? The one introducing each exchange? I found it entertaining and in some cases helpful and in others misleading.	6/5/2018 8:38 AM
5	good to get overview of programme	6/5/2018 8:36 AM
6	It was a nice video, yet not sure whether we should spend time/resources on this, honestly	6/5/2018 8:36 AM
7	It was very useful from a participants point of view. From a conveners point of view, I would do the video again but it is not an easy job, as not everybody is a natural in from of a camera. We need some training and practice on how to do this sort of videos. Otherwise, it can also give the wrong impression to participants.	6/5/2018 8:19 AM
8	The video was just excellent, very helpful to pitch the different EPE sessions on offer. Whether this will be (resource-wise) sustainable for the next EPEs to come is a different matter. I think what made the difference was FAO (and the other Rome-based agencies) communication capabilities to support the preparation of the video (and of all the other EPE materials).	5/28/2018 1:40 PM
9	Very nice way of presenting the sessions	5/28/2018 11:50 AM
10	Maybe not useful but a nice effort. It might have been more useful if it was sent out earlier - before the opening. If it was and I missed it, sorry.	5/25/2018 10:01 AM

11	nice format, engaging, helped deciding which session to attend. also helped connecting faces with names.	5/25/2018 9:49 AM
12	It gave me adequate information to decide which sessions I would attend	5/24/2018 10:12 PM
13	Maybe doing a bit shorter videos. But it did give a good impression of what would happen	5/24/2018 8:20 AM
14	Well done!	5/23/2018 9:44 PM
15	Well done. Key messages were clear and straightforward.	5/23/2018 6:49 PM
16	I chose my sessions based on the agenda on the conference booklet: 1. The title of the session- Did it catch my attention? Was it relevant for me? 2. The description of the session. I think the video might be confounding, because some people are perhaps better speakers or "sellers" of their session, and might encourage you to come to their session, instead to a session that is more relevant but wasn't "sold" as well in the video.	5/23/2018 5:22 PM
17	The video itself was a bit long and hard to follow.	5/23/2018 4:16 PM

Q20 As a participant, were you satisfied with the overall preparation of the EPE (dissemination of information, clarity of content, etc.)?

Answered: 64 Skipped: 9



ANSWER CHOICES	RESPONSES	
Very satisfied	68.75%	44
Satisfied	31.25%	20
Dissatisfied	0.00%	0
Very dissatisfied	0.00%	0
TOTAL		64

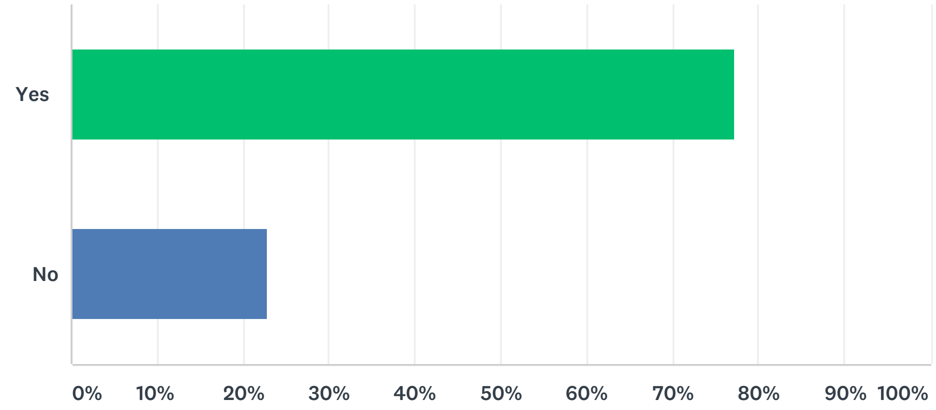
Q21 Please write down any comment about the preparation phase of the EPE as participant, lead convener or collaborator

#	RESPONSES	DATE
1	It was satisfying to collaborate with other UN agency colleagues!	6/5/2018 5:50 PM
2	It is crucial that the EPE emphasizes how is evaluation (the profession and the United Nations) position within a wider context vis a vis a (public policy, development agenda, post truth, audit, big data, complexity etc, etc). Looking forward ten years from now (the use and the implication of articial intelligence for our profession, etc)	6/5/2018 3:07 PM
3	I attended one EPE that attracted a larger group so the group was divided. Although the subject matter was highly relevant to my work, the presenters didn't stay on topic well so the time wasn't well spent in the end. And the keynote contributor to the topic wasn't in the group that I attended.	6/5/2018 11:22 AM
4	All fine	6/5/2018 11:13 AM
5	The EPE was very well organized. It seemed everybody in the organization made signifficant efforts to improve and raise the bar of the EPE	6/5/2018 10:02 AM
6	The time allocated to the session did not allow to reach the depth in the discussion. Maybe a bit longer session would give this possibility.	6/5/2018 9:20 AM
7	Accuracy of info could have been better. I had scheduled aperitivo on Wednesday based on first agenda so an updated agenda could have been shared earlier.	6/5/2018 8:38 AM
8	The precise agenda came in a bit late. All the rest superb!	6/5/2018 8:36 AM
9	The whole EPE preparation phase was excellent. The organizers did facilitate the process in a very professional manner.	6/5/2018 8:19 AM
10	None	6/4/2018 10:48 PM
11	organized	5/28/2018 3:33 PM
12	See earlier comment.	5/28/2018 1:40 PM
13	excellent preparation. Many thanks to the entire team!	5/25/2018 9:49 AM
14	The EPE this year was great, well designed , useful and engaging	5/25/2018 6:44 AM
15	The videos helped and I appreciate the fact that the agenda was sent in advance. Next time, for the sake of new comers, it may be good to share participant's list	5/24/2018 10:12 PM
16	I received the agenda of the EPE a bit too late (few days before the event) but I really appreciated the survey to indicate in what session we could possibly participate	5/24/2018 10:47 AM
17	I enjoyed the interactive, innovative way all these EPEs worked.	5/24/2018 8:20 AM
18	no comment	5/23/2018 10:20 PM
19	As EPE participant, I valued the knowledge exchange and safe space for discussion that was created.	5/23/2018 9:44 PM
20	All went very well	5/23/2018 6:49 PM
21	More preparatory meetings required. At least one more	5/23/2018 5:29 PM
22	It would be useful to have co-conveners, as our convener was not able to attend on time	5/23/2018 5:08 PM
23	we may want to consider involving more staff at regional and country level	5/23/2018 4:47 PM
24	the sessions I participated in where relevant and inclusive. I thought that the EPE managed to be what it is supposed to be an exchange between practioners.	5/23/2018 4:45 PM
25	The preparation was very well done and impressive. Thank you very much for your hard work.	5/23/2018 4:16 PM
26	None	5/23/2018 4:13 PM

27	very good	5/23/2018 4:13 PM
28	Every member of the EPE committee was brilliant!! ;)	5/23/2018 6:15 AM
29	The role of “lead convener” and “collaborator” may not have been very clear, especially at the stage of call for interest. Perhaps there was some lack of clarity in whether “collaborators” were those who thought who had something to say/share on the topic (this was the case at the time of call for interest - but some also proposed the modality) or who were to help and work with the lead conveners (esp. facilitation of group discussion - this is how it was). Perhaps it would be better to have “lead conveners’ (or lead co-conveners) on certain topics first, and then seek collaborators who are interested in the same topic and in working with lead conveners along the proposed modality and content.	5/22/2018 11:07 PM

Q22 Did you read/use the Logistical note sent by the EPE Organization Committee?

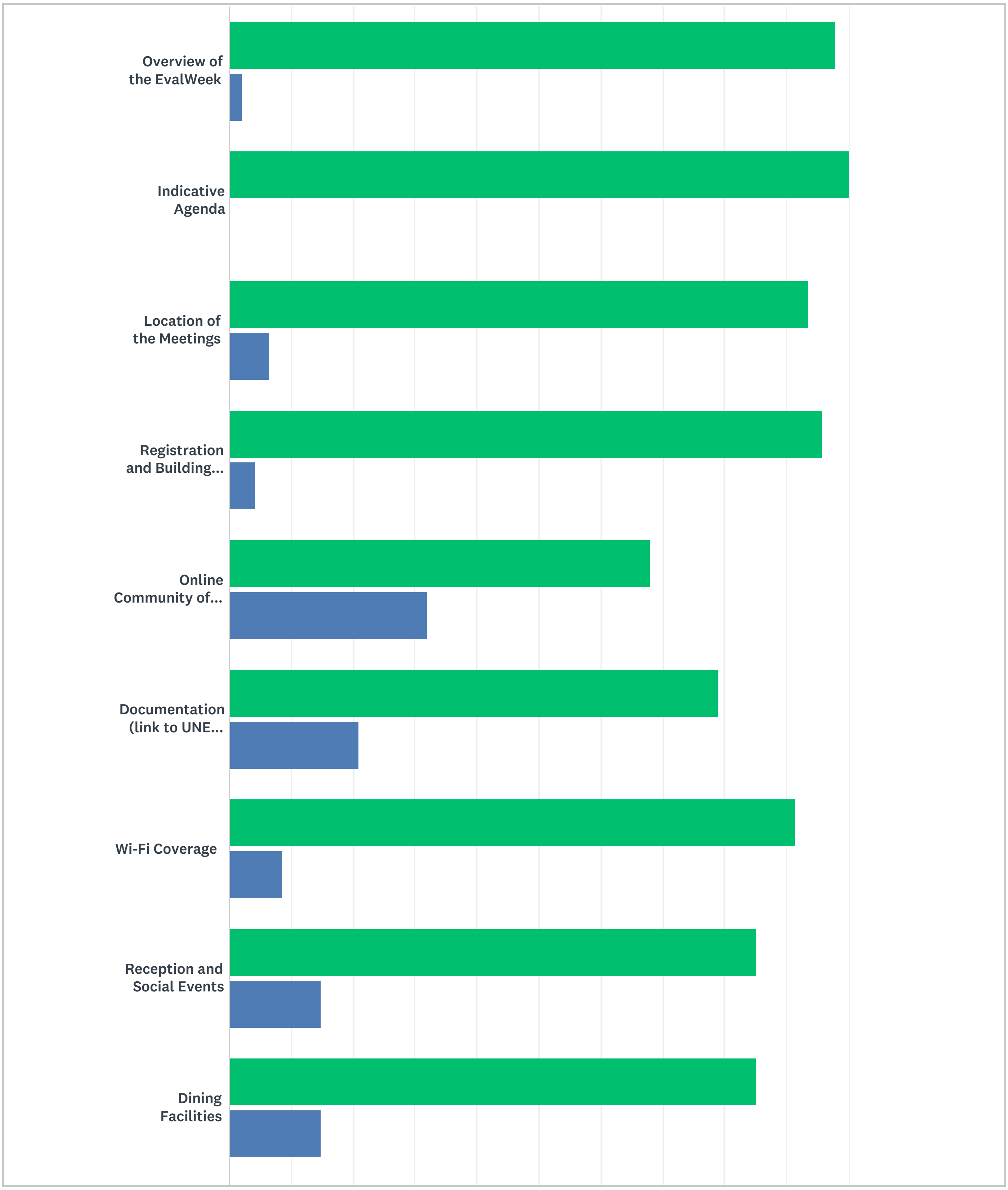
Answered: 66 Skipped: 7

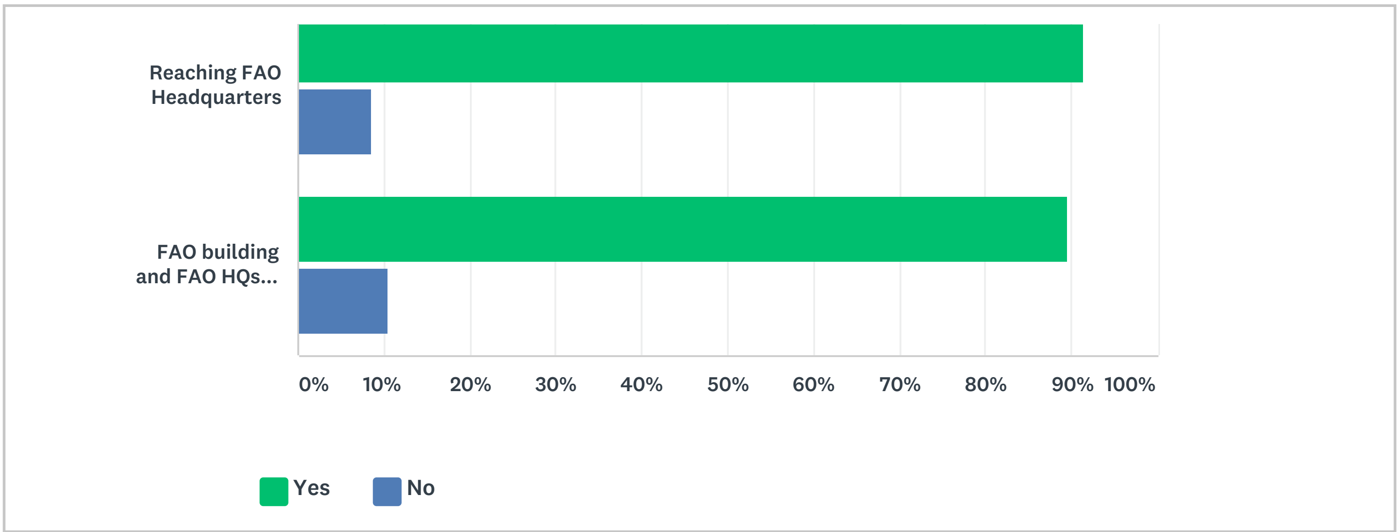


ANSWER CHOICES	RESPONSES	
Yes	77.27%	51
No	22.73%	15
Total Respondents: 66		

Q23 In the first part of the Logistical note (“meeting information”), were the following sections useful?

Answered: 49 Skipped: 24





	YES	NO	TOTAL
Overview of the EvalWeek	97.87% 46	2.13% 1	47
Indicative Agenda	100.00% 48	0.00% 0	48
Location of the Meetings	93.62% 44	6.38% 3	47
Registration and Building Access	95.74% 45	4.26% 2	47
Online Community of Practice	68.09% 32	31.91% 15	47
Documentation (link to UNEG website)	79.17% 38	20.83% 10	48
Wi-Fi Coverage	91.49% 43	8.51% 4	47
Reception and Social Events	85.11% 40	14.89% 7	47
Dining Facilities	85.11% 40	14.89% 7	47
Reaching FAO Headquarters	91.49% 43	8.51% 4	47
FAO building and FAO HQs Orientation Map	89.58% 43	10.42% 5	48

#	ANY COMMENT ON THE “MEETING INFORMATION” SECTION OF THE LOGISTICAL NOTE?	DATE
1	Pictures are not needed	6/5/2018 11:13 AM
2	One social event was canceled. Would have been nice to save some tables in cafeteria for participants	6/5/2018 8:41 AM
3	Well prepared	5/28/2018 3:34 PM
4	Well done	5/23/2018 6:50 PM

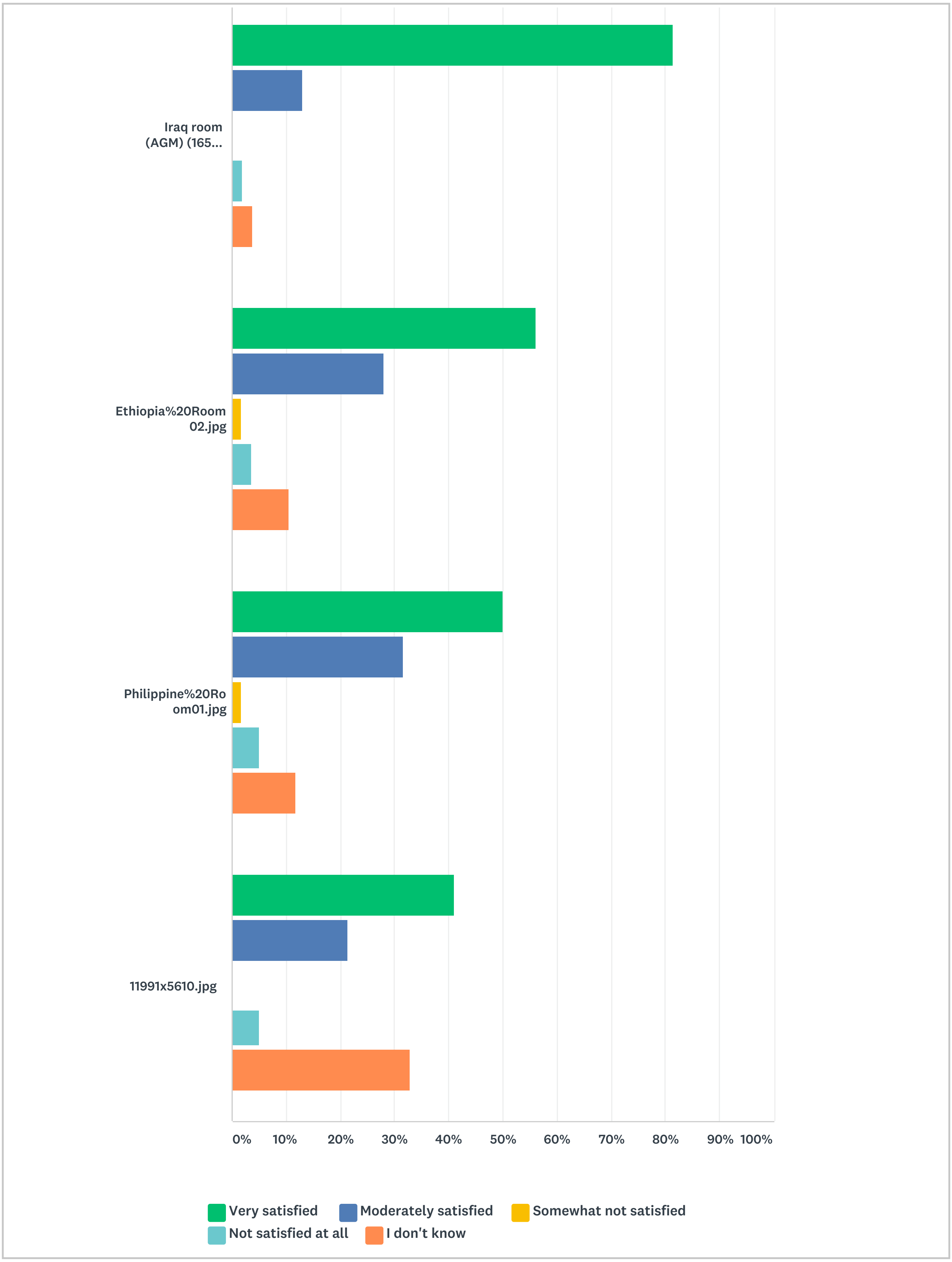
Q24 Any comment on the “services available at FAO” section of the logistical note? Was anything missing?

Answered: 14 Skipped: 59

#	RESPONSES	DATE
1	No	6/5/2018 5:51 PM
2	All fine	6/5/2018 11:13 AM
3	don’t recall	6/5/2018 8:37 AM
4	Every thing was perfect.	6/5/2018 8:19 AM
5	None	6/4/2018 10:49 PM
6	No well done	5/30/2018 7:46 PM
7	All good	5/28/2018 3:34 PM
8	I think the organizers did a great job of providing guidance so once at the event I didn’t need the note	5/25/2018 10:02 AM
9	no comment	5/23/2018 10:21 PM
10	No	5/23/2018 6:50 PM
11	no	5/23/2018 5:30 PM
12	All good	5/23/2018 4:33 PM
13	Perhaps local services around FAO could also be indicated for use during free time.	5/23/2018 4:17 PM
14	None	5/23/2018 4:14 PM

Q25 Were you satisfied with the choice and quality of the meeting rooms for the PDS, EPE and AGM?

Answered: 61 Skipped: 12

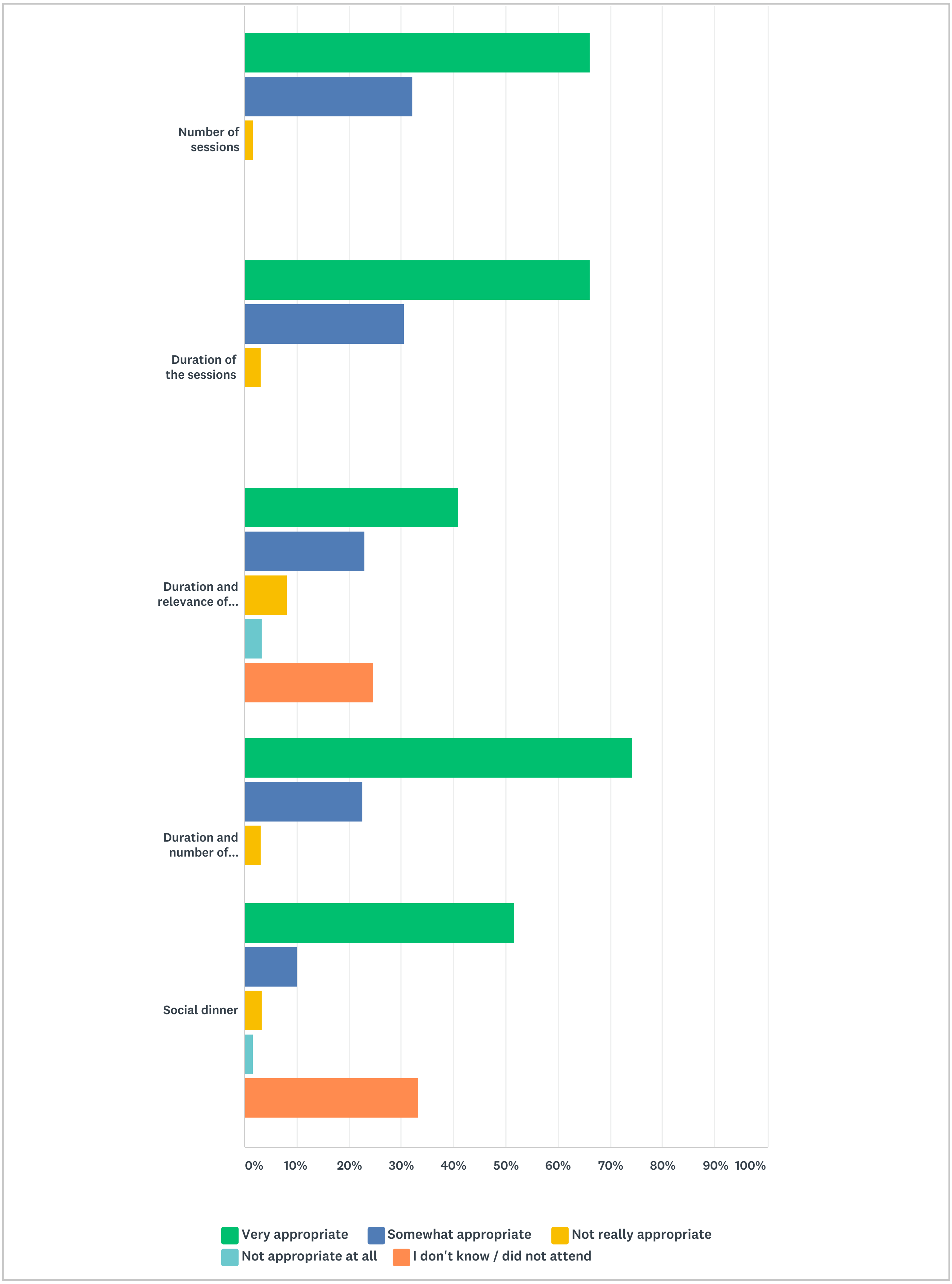


	VERY SATISFIED	MODERATELY SATISFIED	SOMEWHAT NOT SATISFIED	NOT SATISFIED AT ALL	I DON'T KNOW	TOTAL
Iraq room (AGM) (165 people)	81.48% 44	12.96% 7	0.00% 0	1.85% 1	3.70% 2	54
	56.14% 32	28.07% 16	1.75% 1	3.51% 2	10.53% 6	57
	50.00% 30	31.67% 19	1.67% 1	5.00% 3	11.67% 7	60
	40.98% 25	21.31% 13	0.00% 0	4.92% 3	32.79% 20	61

#	COMMENTS ON THE MEETING ROOMS. WHICH ROOM DID YOU LIKE THE MOST AND WHY?	DATE
1	Ethiopia room was lovely but not conducive to group work.	6/5/2018 5:51 PM
2	Ethiopia and Nigeria a bit small for group work	6/5/2018 3:08 PM
3	Rooms with windows would be preferrable. I liked the Iraq room the best	6/5/2018 8:39 AM
4	Rooms with windows	6/4/2018 10:50 PM
5	The Philippines room was too cold (AC)	5/30/2018 7:48 PM
6	Iraq, ethiopia and philippines room were great. Iraq room allowed for group work given the table setting - which was absolutely needed!	5/25/2018 9:52 AM
7	Ethiopia. Very spacious.	5/24/2018 10:13 PM
8	Interactive sessions would rather need rooms where tables can be moved around like in the Iraq room	5/24/2018 8:22 AM
9	Nigeria was too small for group activities	5/23/2018 5:30 PM
10	Philippines room had weird acoustics. I couldnt always hear well when sitting in the rear	5/23/2018 5:24 PM
11	The rooms were not suitable for the kind of group work/interaction/ world cafe styles foreseen in EPEs	5/23/2018 5:10 PM
12	Difficult to have group discussions in the Ethiopia room. It was too noisy.	5/23/2018 2:59 PM
13	The Ethiopia room was beautiful but not very conducive to group work	5/23/2018 6:27 AM
14	It would have been good to have meeting rooms with more flexibility, e.g. to allow round table set-up to facilitate group discussions.	5/22/2018 11:13 PM

Q26 Please comment about the following aspects related to the organization of the EPE

Answered: 62   Skipped: 11

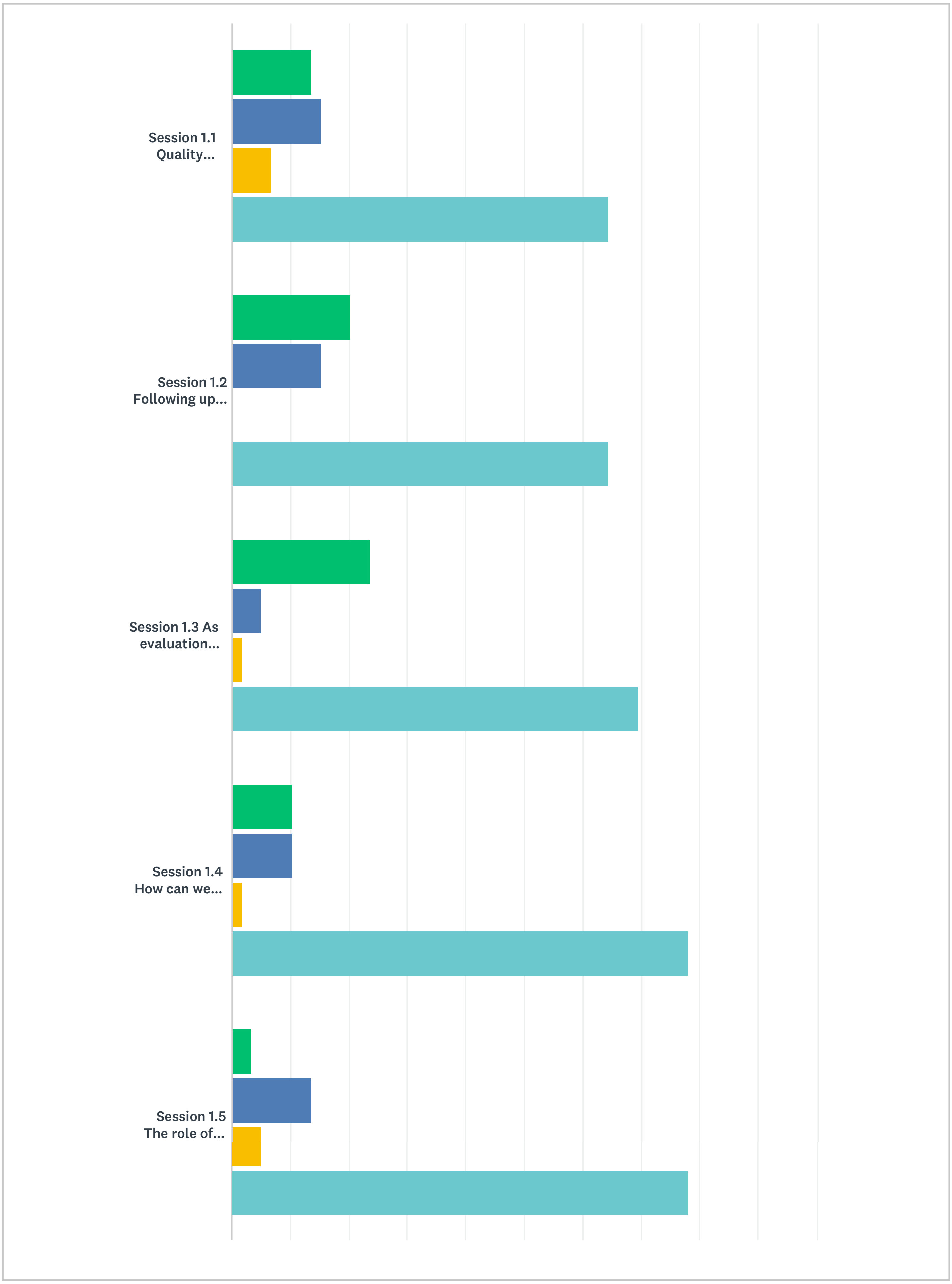


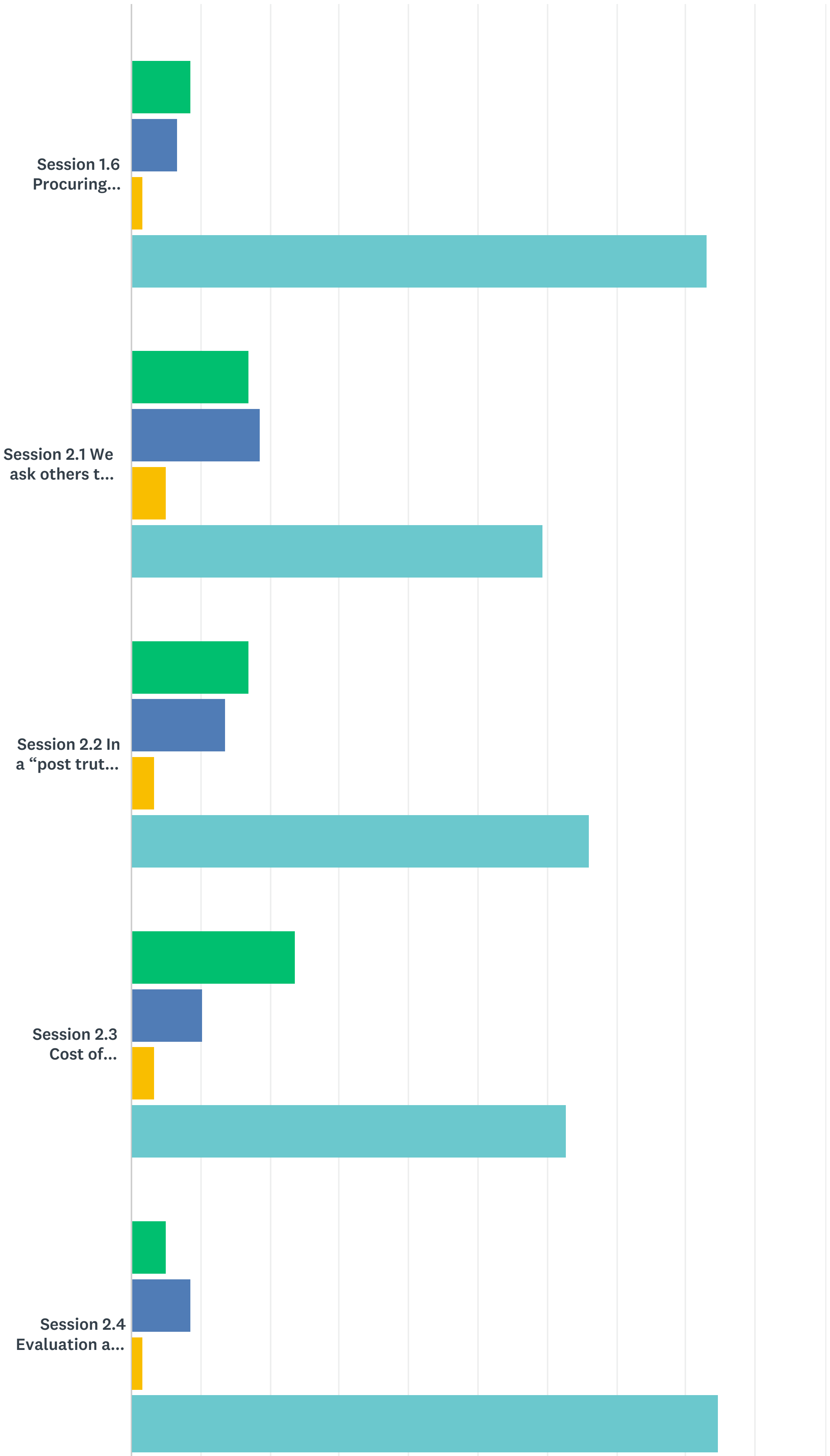
	VERY APPROPRIATE	APPROPRIATE SOMEWHAT	NOT REALLY APPROPRIATE	NOT APPROPRIATE AT ALL	I DON'T KNOW / DID NOT ATTEND	TOTAL
Number of sessions	66.13% 41	32.26% 20	1.61% 1	0.00% 0	0.00% 0	62
Duration of the sessions	66.13% 41	30.65% 19	3.23% 2	0.00% 0	0.00% 0	62
Duration and relevance of lunch time seminars	40.98% 25	22.95% 14	8.20% 5	3.28% 2	24.59% 15	61
Duration and number of coffee breaks	74.19% 46	22.58% 14	3.23% 2	0.00% 0	0.00% 0	62
Social dinner	51.67% 31	10.00% 6	3.33% 2	1.67% 1	33.33% 20	60

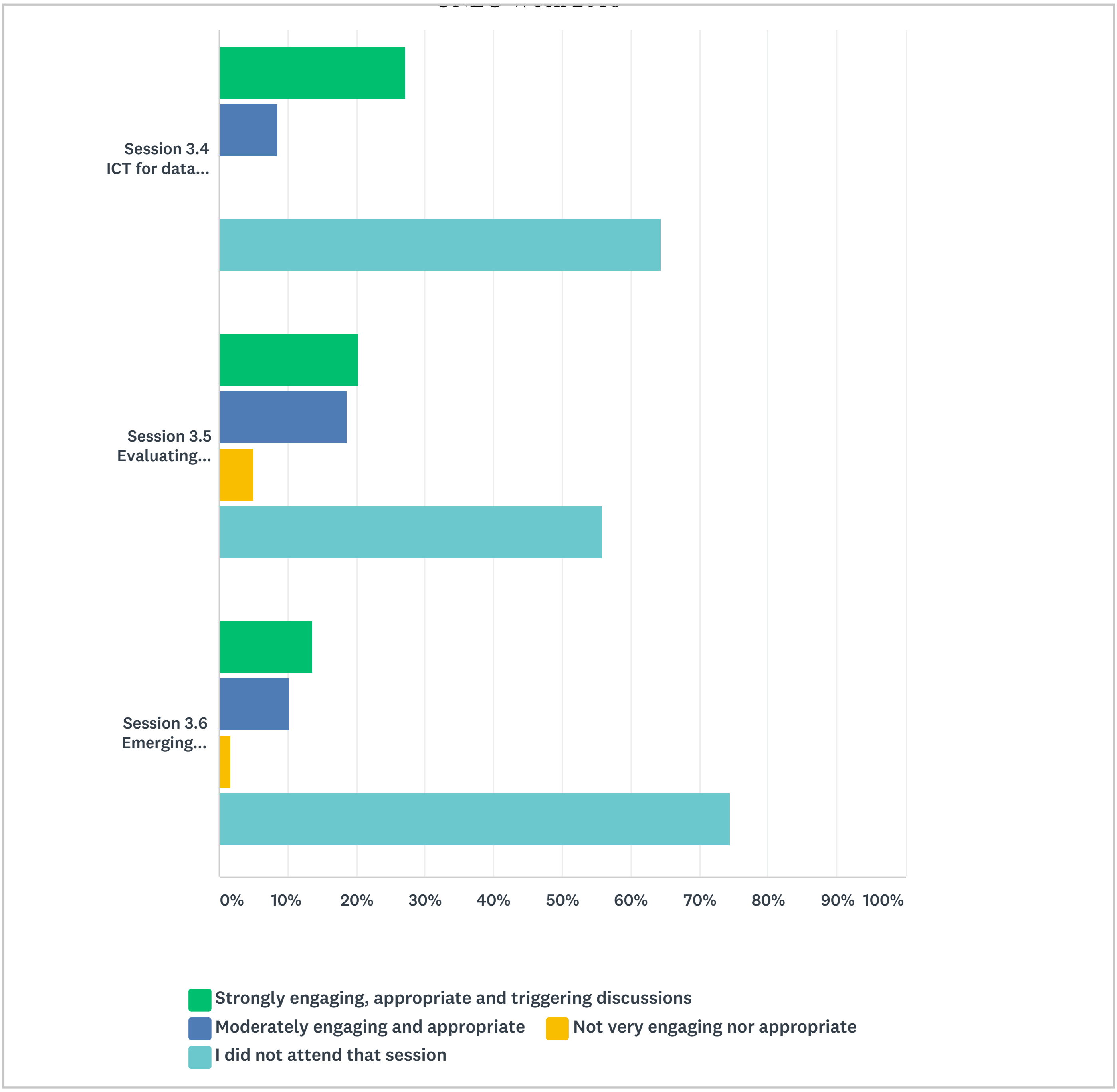
#	PLEASE GIVE US YOUR GENERAL COMMENTS OF THE AGENDA AND ORGANIZATION OF THE TIME DURING THE EPE	DATE
1	There were too many sessions, while there was not that much content to be shared. I suggest to reduce the length of the EPE and increase instead the duration of technical masterclass bringing specialists from different angles to provide teaching out of the box materials. Sharing experiences is interesting to a moderate extent. Once we are aware of the experiences of others, there is little else to share.	6/5/2018 10:06 AM
2	There were many people who came for the dinner. However, the set up and arrangements of the tables allowed to speak to only few of them during the whole dinner.	6/5/2018 9:23 AM
3	Lunch together was a missed opportunity	6/5/2018 8:43 AM
4	I believe there should be both sessions of 1.5-2 hours and longer ones, of 4 hours, as some topics warrant more indepth learning. Lunchbreak sessions I found too short	6/5/2018 8:39 AM
5	I only have praises for the excellent work of the organizing committee. Well done!	6/5/2018 8:20 AM
6	None	6/4/2018 10:50 PM
7	The programme was already quite tight to also attend the lunch time seminars	6/4/2018 5:53 PM
8	Due to lunch time, poor attendance on those seminars	5/29/2018 5:40 PM
9	Even less EPE sessions would have worked just fine for the programme. I understand there are also room capability issues to be considered.	5/28/2018 1:43 PM
10	lunch time seminars and videos shown sometimes difficult to attend due to preparation for workshops or time needed to move from room to room.	5/25/2018 9:52 AM
11	The dinner was a great idea to allow people to socialize outside of work. I enjoyed it.	5/24/2018 10:13 PM
12	Well done! Good combo of short sessions and enough pauses in between	5/24/2018 8:22 AM
13	Excellent work	5/23/2018 5:30 PM
14	It would have been useful to have some more time to discuss bilaterally/ in smaller groups, as the sessions were very much back to back, including lunch seminars	5/23/2018 5:10 PM
15	some of the lunch time seminars could have been fully fledged sessions. for example the one on decentralized evaluation function. this is a subject that deserves more attention and careful analysis.	5/23/2018 4:50 PM
16	could have less sessions and extend the length of some of the topics and reduce the length of others according to the amount of content to discuss.	5/23/2018 4:19 PM

Q27 Please comment about the modality of the EPE sessions you attended

Answered: 59   Skipped: 14




Session 1.6
Procuring...
Session 2.1 We
ask others t...
Session 2.2 In
a “post trut...
Session 2.3
Cost of...
Session 2.4
Evaluation a...



	STRONGLY ENGAGING, APPROPRIATE AND TRIGGERING DISCUSSIONS	MODERATELY ENGAGING AND APPROPRIATE	NOT VERY ENGAGING NOR APPROPRIATE	I DID NOT ATTEND THAT	TOTAL
Session 1.1 Quality Assurance beyond the checklist: how can we ensure the reports are evidence-based?	13.56% 8	15.25% 9	6.78% 4	64.41% 38	59
Session 1.2 Following up on recommendations: what are the successful practices?	20.34% 12	15.25% 9	0.00% 0	64.41% 38	59
Session 1.3 As evaluation managers, we all face similar challenges: let's share them and learn from them	23.73% 14	5.08% 3	1.69% 1	69.49% 41	59
Session 1.4 How can we ensure a gender focused evaluation?	10.17% 6	10.17% 6	1.69% 1	77.97% 46	59
Session 1.5 The role of donors in the evaluations: how should we involve them and how to manage expectations?	3.39% 2	13.56% 8	5.08% 3	77.97% 46	59

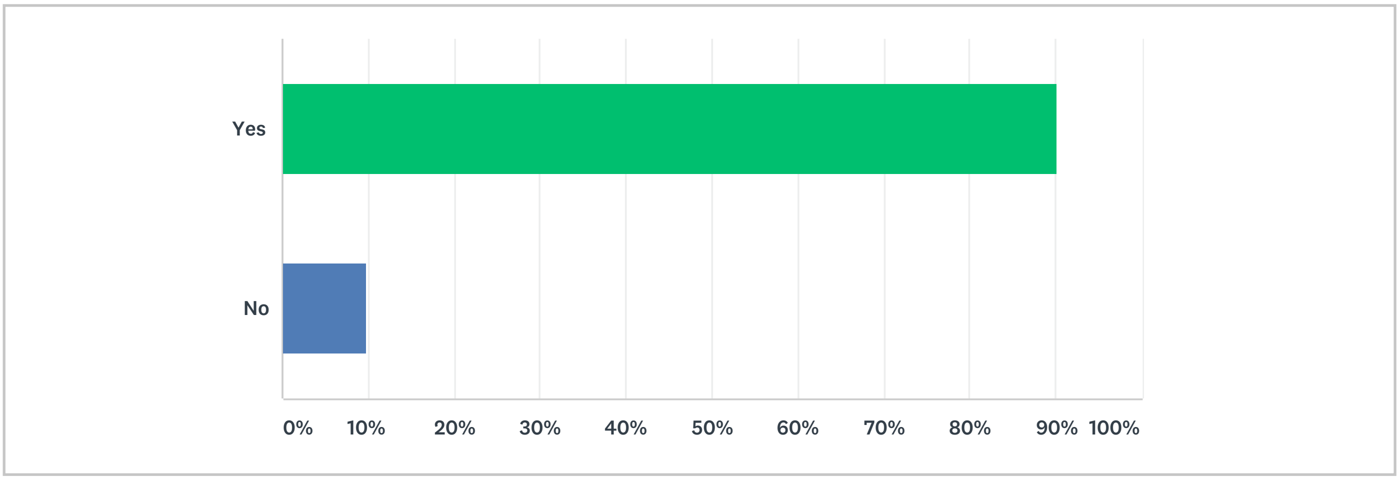
Session 1.6 Procuring evaluations externally: current practices and areas for improvement	8.47% 5	6.78% 4	1.69% 1	83.05% 49	59
Session 2.1 We ask others to demonstrate impact: how can we demonstrate the impact of our evaluations?	16.95% 10	18.64% 11	5.08% 3	59.32% 35	59
Session 2.2 In a "post truth" era, how can we evaluators ensure we adapt and remain heard?	16.95% 10	13.56% 8	3.39% 2	66.10% 39	59
Session 2.3 Cost of evaluations: how can we make efforts to rationalize the use of resources in evaluations? And what are the best practices for low-cost evaluations?	23.73% 14	10.17% 6	3.39% 2	62.71% 37	59
Session 2.4 Evaluation and hindsight: assessing past interventions against yesterday's standards or today's wisdom and criteria?	5.08% 3	8.47% 5	1.69% 1	84.75% 50	59
Session 2.5 What does independence mean today?	28.81% 17	16.95% 10	1.69% 1	52.54% 31	59
Session 2.6 Full Disclosure? What evaluators don't write about	13.56% 8	10.17% 6	1.69% 1	74.58% 44	59
Session 3.1 No baseline: how to measure impact?	13.56% 8	13.56% 8	22.03% 13	50.85% 30	59
Session 3.2 Making the best use of Theories of Change in Evaluations	10.17% 6	8.47% 5	5.08% 3	76.27% 45	59
Session 3.3 Country Programme Evaluations - methodological challenges	8.47% 5	16.95% 10	5.08% 3	69.49% 41	59
Session 3.4 ICT for data analysis in evaluations	27.12% 16	8.47% 5	0.00% 0	64.41% 38	59
Session 3.5 Evaluating Policy support	20.34% 12	18.64% 11	5.08% 3	55.93% 33	59
Session 3.6 Emerging principles and lessons learned for the development of organizational evaluation policies	13.56% 8	10.17% 6	1.69% 1	74.58% 44	59

#	PLEASE PROVIDE YOUR GENERAL COMMENTS ABOUT THE MODALITY OF THE EPE SESSIONS. TELL US WHICH MODALITY(IES) YOU PREFERRED AND WHY!	DATE
1	As i indicated earlier, I suggest to revisit the content of the EPE. My perspective is that after several EPEs, there is little added value in repeatingly sharing experiences over similar subjects. This multiply by x paralleled sessions every 90 minutes during 2 days may be too much. I suggest instead to focus on providing masterclasses on new developments in related areas bringing specialists from outside the UNEG to teach on specific subjects, developments or apps or whatever new.	6/5/2018 10:15 AM
2	On cost session, it was very engaging but not appropriate to our situation. So not what I was expecting or useful to my situation. Perhaps at a higher level it is.	6/5/2018 8:56 AM
3	The modalities were excellent. It was great to move from the traditional panel presentations to more interactive sessions. There was so much cross pollination and I took a lot of very good ideas with me. The only session I attended which was terrible was the one on no baseline: how to measure impact. This was badly prepare the presenter was textually reading from a whole paper. She did not have a power point but she was reading her paper to us. The other presenter from WFP did not appear and no one new where she was. However, I think this things can also happen and surely it was an exception.	6/5/2018 8:38 AM
4	Interesting sessions, no preferred modalities	6/4/2018 11:13 AM

5	Modalities were good - varied and interesting. Oxford style debate worked well. Quality of the contents of the sessions varied a lot. Still rather too insular and UN focused, need to draw a lot more on what's going on outside the UN and involve non-UN people - it will up the quality. That said, it was a very well organized and innovative EPE - so well done. Onward and upward.. :)	5/31/2018 5:44 PM
6	Excellent modalities chosen for the EPE. This was really a different but excellent and well done EPE. Congratulations!	5/26/2018 2:10 AM
7	world café format as well as the hot-chair discussion format were most engaging and inclusive and fun!	5/25/2018 10:05 AM
8	They were great. The switching of tables for session 3.4 'ICT for data analysis in evaluations' was too little time for meaningful discussion. That said, the session was my best but I wish I had more time to ask the facilitator questions.	5/24/2018 10:18 PM
9	The session on independence and full disclosure were definitely my highlights.	5/24/2018 8:29 AM
10	role playing, world cafe sessions were more engaging.	5/23/2018 5:15 PM
11	I also participated in a session on TOC which was good although I think it was not really necessary to have such a long introduction on what a TOC is with three different powerpoints.	5/23/2018 4:53 PM
12	Session 3.1 No baseline: how to measure impact? I was greatly looking forward to this session and was very disappointed. This session was poorly organized and quite frankly did not capture the group's interest. Also the lead convener did not attend due to "logistic reason" which is unacceptable.	5/23/2018 4:26 PM

Q28 Would you like to keep the “no-Power Point” rule for the next EPE?

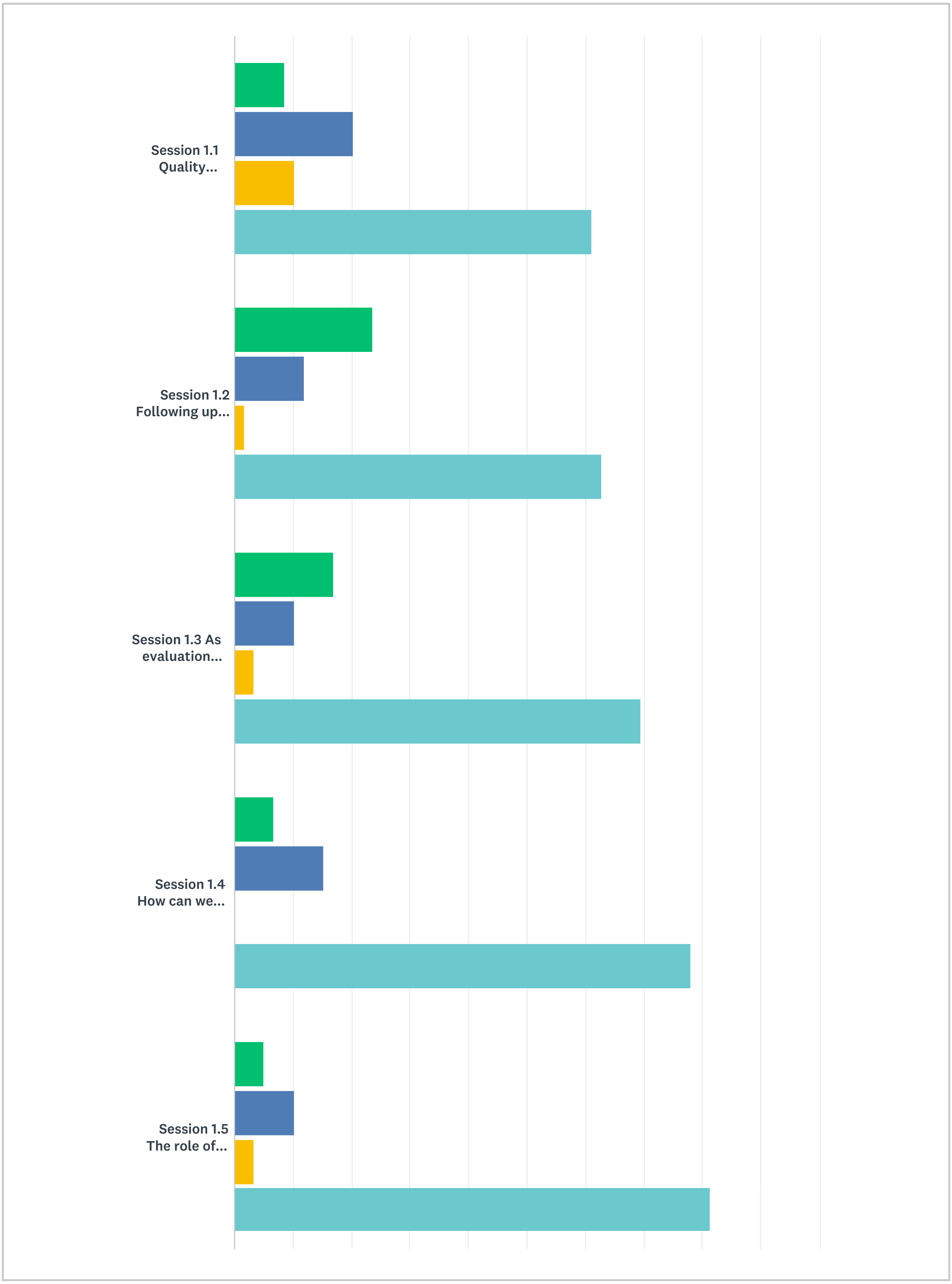
Answered: 51   Skipped: 22

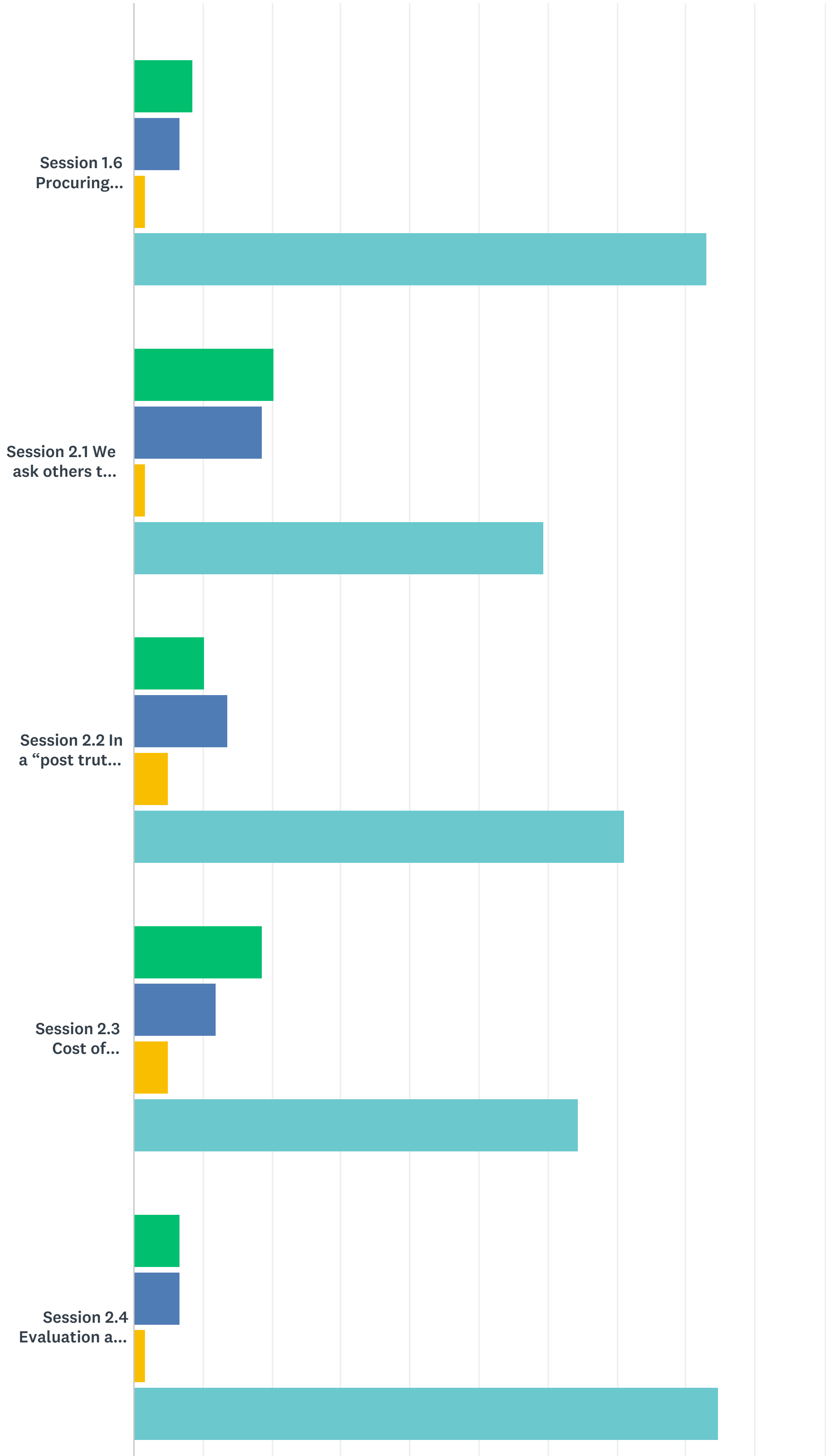


ANSWER CHOICES	RESPONSES	
Yes	90.20%	46
No	9.80%	5
TOTAL		51

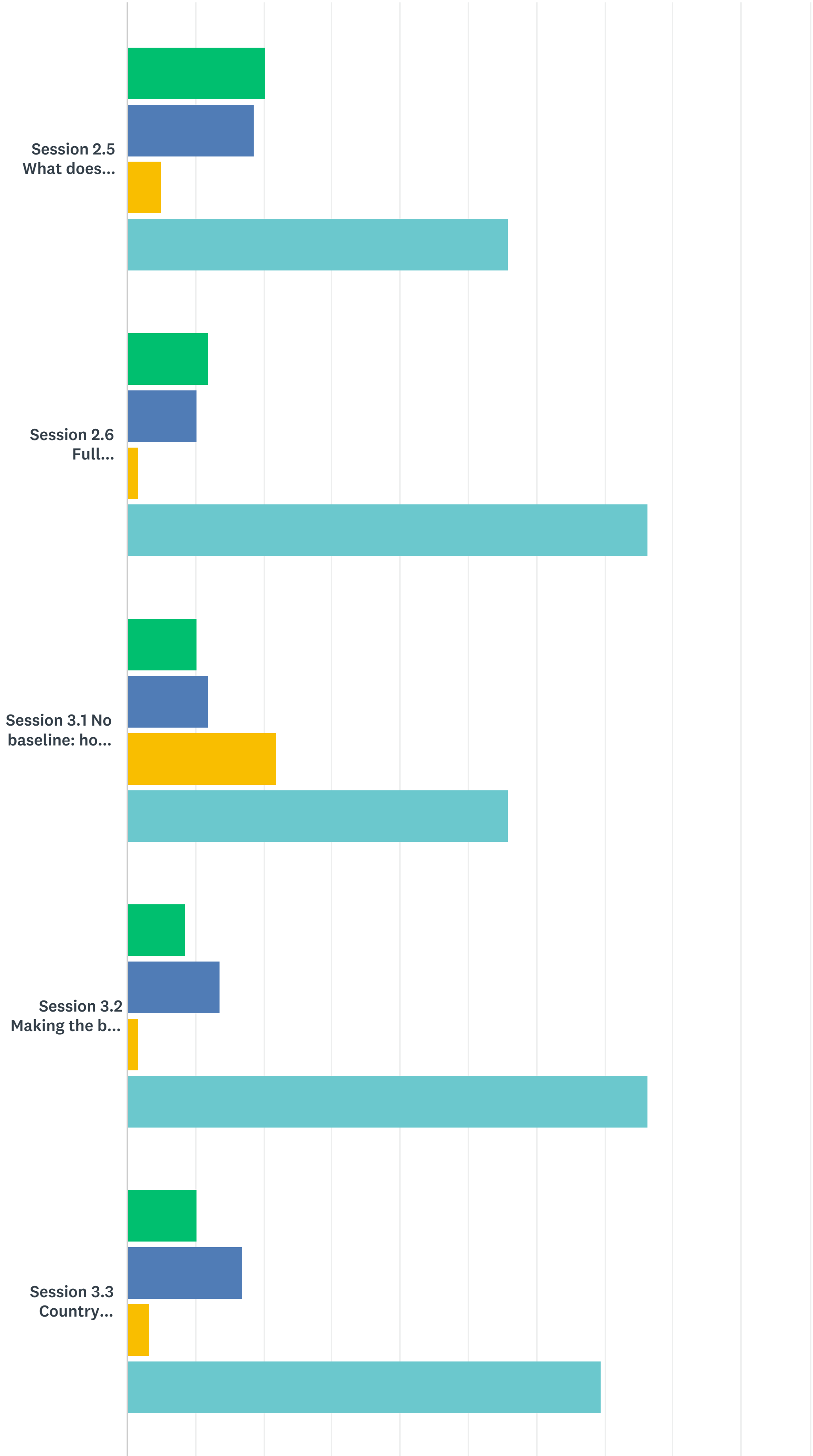
Q29 Please comment about the usefulness of the sessions you attended

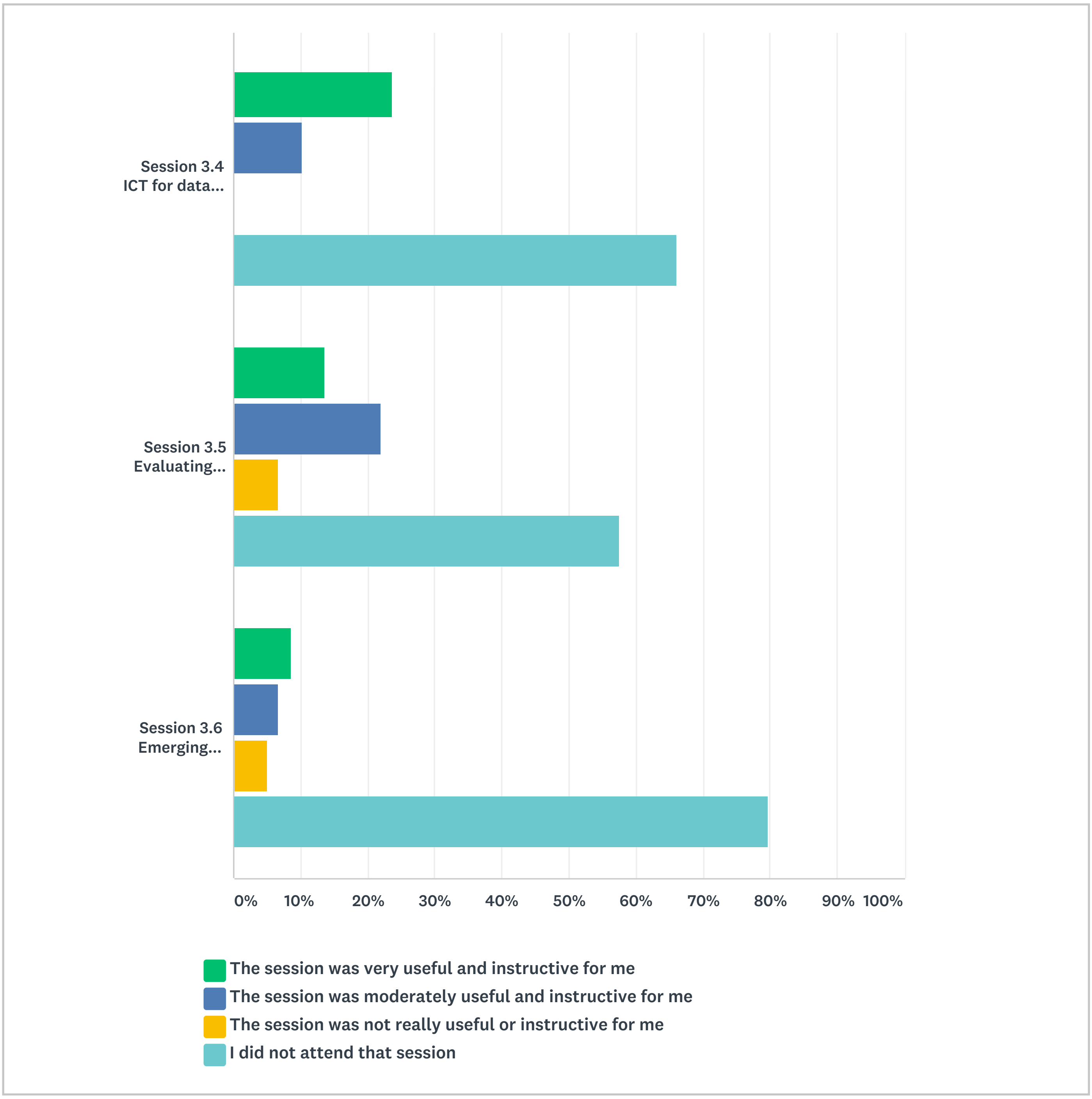
Answered: 59 Skipped: 14





Session 1.6 Procuring...
Session 2.1 We ask others t...
Session 2.2 In a “post trut...
Session 2.3 Cost of...
Session 2.4 Evaluation a...


Session 2.5 What does...
Session 2.6 Full...
Session 3.1 No baseline: ho...
Session 3.2 Making the b...
Session 3.3 Country...



	THE SESSION WAS VERY USEFUL AND INSTRUCTIVE FOR ME	THE SESSION WAS MODERATELY USEFUL AND INSTRUCTIVE FOR ME	THE SESSION WAS NOT REALLY USEFUL OR INSTRUCTIVE FOR ME	I DID NOT ATTEND THAT SESSION	TOTAL
Session 1.1 Quality Assurance beyond the checklist: how can we ensure the reports are evidence-based?	8.47% 5	20.34% 12	10.17% 6	61.02% 36	59
Session 1.2 Following up on recommendations: what are the successful practices?	23.73% 14	11.86% 7	1.69% 1	62.71% 37	59
Session 1.3 As evaluation managers, we all face similar challenges: let's share them and learn from them	16.95% 10	10.17% 6	3.39% 2	69.49% 41	59
Session 1.4 How can we ensure a gender focused evaluation?	6.78% 4	15.25% 9	0.00% 0	77.97% 46	59

Session 1.5 The role of donors in evaluations: how should we involve them and how to manage expectations?	5.08% 3	10.17% 6	3.39% 2	81.36% 48	59
Session 1.6 Procuring evaluations externally: current practices and areas for improvement	8.47% 5	6.78% 4	1.69% 1	83.05% 49	59
Session 2.1 We ask others to demonstrate impact: how can we demonstrate the impact of our evaluations?	20.34% 12	18.64% 11	1.69% 1	59.32% 35	59
Session 2.2 In a "post truth" era, how can we evaluators ensure we adapt and remain heard?	10.17% 6	13.56% 8	5.08% 3	71.19% 42	59
Session 2.3 Cost of evaluations: how can we make efforts to rationalize the use of resources in evaluations? And what are the best practices for low-cost evaluations?	18.64% 11	11.86% 7	5.08% 3	64.41% 38	59
Session 2.4 Evaluation and hindsight: assessing past interventions against yesterday's standards or today's wisdom and criteria?	6.78% 4	6.78% 4	1.69% 1	84.75% 50	59
Session 2.5 What does independence mean today?	20.34% 12	18.64% 11	5.08% 3	55.93% 33	59
Session 2.6 Full Disclosure? What evaluators don't write about	11.86% 7	10.17% 6	1.69% 1	76.27% 45	59
Session 3.1 No baseline: how to measure impact?	10.17% 6	11.86% 7	22.03% 13	55.93% 33	59
Session 3.2 Making the best use of Theories of Change in Evaluations	8.47% 5	13.56% 8	1.69% 1	76.27% 45	59
Session 3.3 Country Programme Evaluations - methodological challenges	10.17% 6	16.95% 10	3.39% 2	69.49% 41	59
Session 3.4 ICT for data analysis in evaluations	23.73% 14	10.17% 6	0.00% 0	66.10% 39	59
Session 3.5 Evaluating Policy support	13.56% 8	22.03% 13	6.78% 4	57.63% 34	59
Session 3.6 Emerging principles and lessons learned for the development of organizational evaluation policies	8.47% 5	6.78% 4	5.08% 3	79.66% 47	59

#	PLEASE PROVIDE YOUR GENERAL COMMENTS ABOUT THE CONTENT AND RELEVANCE OF THE EPE SESSIONS. AND TELL US WHICH TOPIC(S) YOU PREFERRED!	DATE
1	Cost session was misleading as I thought it was more practical. The discussion of cost was centered around FAO's business model. A fun cathartic session but not useful	6/5/2018 8:56 AM
2	Cost reduction was my preferred one	6/5/2018 8:43 AM
3	Actually, all the topics were very relevant and added a lot of value to our work, in fact, I had difficulties choosing them because I would have like to attend more topics than the ones I did. very much would like to expand on evaluating policy support and look at complexity methodologies and system thinking. I also found the session 1.6 with questions and answers very good, it made me think out of the box.	6/5/2018 8:38 AM
4	Rated 'moderately useful' for most sessions given that I know the topics or the problems debated but interesting debates and good opportunities to refresh some notions. No preferred topics as many were interesting, even difficult to select sometimes. Some sessions were more interactive than others.	6/4/2018 11:13 AM

5	As per my comments in previous question, the quality of the content didn’t match the innovativeness of the modality. This can be built on for next year, but should involve outsiders who may be more involved in the content of the work - and can bring more insight to the topics.	5/31/2018 5:44 PM
6	Overall very relevant sessions. Content was also good	5/26/2018 2:10 AM
7		5/24/2018 1:59 PM
8	The survey comes a bit late. I didn’t remember all the sessions I attended.	5/23/2018 5:29 PM
9	Session 3.1 No baseline: how to measure impact? Since this session was poorly organized, it was not very useful at all. The topic is very interesting but should have been presented better.	5/23/2018 4:26 PM

Q30 Comments or remarks about your experience of the EPE as a participant, lead convener or collaborator

Answered: 23 Skipped: 50

#	RESPONSES	DATE
1	It was very engaging!	6/5/2018 5:53 PM
2	I encourage a more open and participatory approach to all UN evaluation community at all levels and from all organisations with special attention to the young people entering the field of evaluation in the UN	6/5/2018 3:18 PM
3	As mentioned earlier. I think the current principle (sharing experiences) adds little value	6/5/2018 10:15 AM
4	Very engaging	6/5/2018 8:56 AM
5	Very useful! Short and sharp. Created unique team atmosphere	6/5/2018 8:43 AM
6	Dear organizers, thank you very much for this great professional experience. To me it was unique. I have never been in such well prepare meeting and I have been in a lot of meetings. This EPE was best ever.	6/5/2018 8:38 AM
7	excellent, best ever	6/4/2018 6:15 PM
8	Great efforts from colleagues to propose very interesting EPE sessions and diversified. Good work!	6/4/2018 11:13 AM
9	Good turn out, good panellists and interesting debate.	5/31/2018 5:44 PM
10	It was a very good opportunity to learn and share experience. It was well organized.	5/28/2018 3:48 PM
11	I participate as collaborator and participant. I am very satisfied with the experience.	5/26/2018 2:10 AM
12	It had very relevant topics and I learned a lot.	5/24/2018 10:18 PM
13	The EPE was a great learning and networking experience, even though it became a little too tiring at the end of day 2. I would suggest a different modality for the wrap up session, there were definitely too many sessions to be discussed and the time and energy in the room did not allow for it. I would use the Mentimeter only once or twice over the plenary sessions, otherwise there's a risk of losing it's energizing effect	5/24/2018 10:58 AM
14	I really liked the sessions, and those which were most interactive and well planned were the most useful due to the discussion that happened	5/24/2018 8:29 AM
15	no comment	5/23/2018 10:29 PM
16	The EPE format allow for more knowledge sharing than previous sessions. This was personally fulfilling and took the discussions to a more practical level	5/23/2018 9:48 PM
17	Overall a great event. Great to meet all Evaluators of the different agencies.	5/23/2018 5:29 PM
18	The content and variety of sessions were much better than previous years. Small group works helped me to get a deeper insight to different organization's work and particular challenges.	5/23/2018 5:15 PM
19	Great experience. One small detail that I think made a difference was that titles were removed from the badges breaking down the hierarchy and enabling discussion across P and D grades.	5/23/2018 4:53 PM
20	I greatly enjoyed attending this session. It was a very enlightening way to learn other organization's best practice and meet other evaluators in a learning environment.	5/23/2018 4:26 PM
21	very good	5/23/2018 4:17 PM
22	Nio	5/23/2018 4:16 PM
23	great diversity of sessions and approaches	5/23/2018 6:35 AM

Q31 What can be done to better ensure that the EPE sessions respond to the needs and interests of participants? Is there any other topic you would like to discuss during the next EPE?

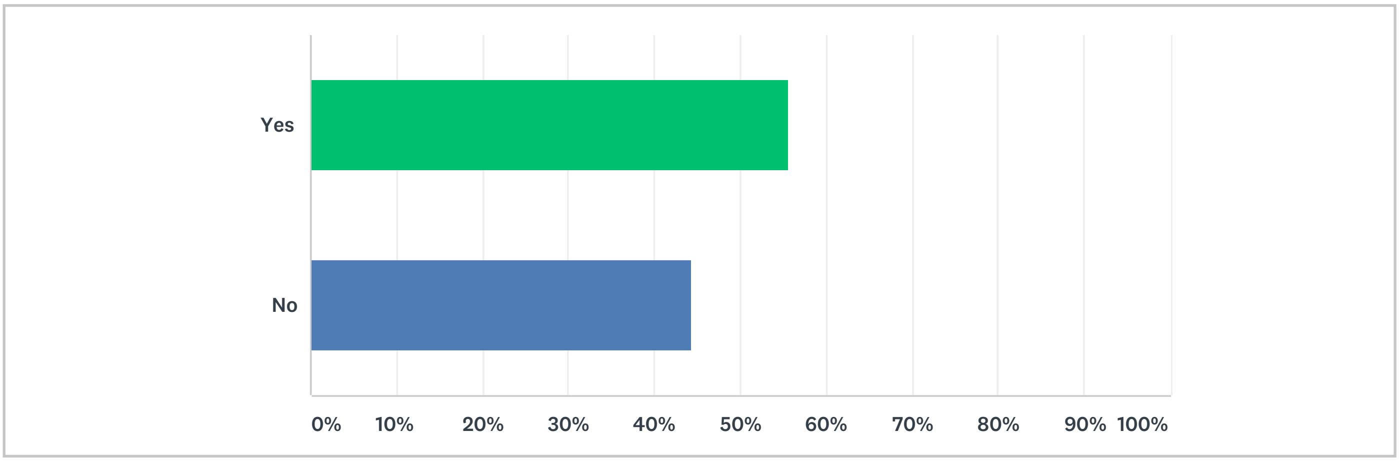
Answered: 24 Skipped: 49

#	RESPONSES	DATE
1	Please make sure you do the same survey you did this year before desigigning the new EPE. 1. How artificial intelligence is going to affect the profession of evaluation 2. How does the fussion between audit and evaluation in an oversight office affect the learning fuction on evaluation ? 3. What is the UN doing to ensure that SDgs are going to be measure ? 4. How can innovation be evaluated? 5. How can data visualizations increase impact of evaluation results?	6/5/2018 3:18 PM
2	I found the PD format of more value than the EPE format. I like the variety of workshops available through the EPE but I would suggest more that EPEs gravitate more to the PDS format and a shorter list of topics to provide stronger outcomes. In case of the PD event I participated in, both formats supported professional exchange and network development.	6/5/2018 11:30 AM
3	I would suggest to perhaps reduce it to a 2 day total (combined with the 1 day of the week), and transform it into a masterclass type of sessions bringing new developers from outside UNEG to teach innovations.	6/5/2018 10:15 AM
4	Lunch sessions need not be paid for but could be organized so that networking can continue.	6/5/2018 8:56 AM
5	More on NECD	6/5/2018 8:43 AM
6	I think it would be difficult for anyone to fit in your shoes (organizers 'shoes). Therefore, it might be good if the organizers to help and provide advice to the next EPE. Evaluation of policies/ normative work. Specially looking into complex theories and system thinking. More on social research methods and facilitation skills approaches which are essential for evaluators. What about leadership skills to influence implementation of recommendations.	6/5/2018 8:38 AM
7	hold over a longer period	6/4/2018 6:15 PM
8	Use of common tools for UNEG members, sharing the costs of development/maintenance. Evaluation database, software...	6/4/2018 5:50 PM
9	Very diversified topics, the needs were certainly well covered. No specific topic in mind that could be missing.	6/4/2018 11:13 AM
10	See earlier comments / personal views.	5/31/2018 5:44 PM
11	Good continue doing need assessment Focus on practical things than theory	5/28/2018 3:48 PM
12	Swift and low cost evaluations - Shall we unquestionably follow this new paradigm?	5/26/2018 2:10 AM
13	more SDG focus - maybe even the structure of the EPE?	5/25/2018 10:05 AM
14	Given the interest showing atvthe EPE evaluating policy support should be widely covered at the next EPE and PDS	5/25/2018 6:53 AM
15	Desk rotations, when used, should be longer than 15 minutes.	5/24/2018 10:18 PM
16	Keep the interactivity! There are plenty of ways to facilitate sessions like these, maybe thus engaging people who know how to organize such events? For the next time, add "How can we use innovative approaches to evaluations in rather risk-averse organizations"	5/24/2018 8:29 AM
17	no comment	5/23/2018 10:29 PM
18	Do a pre-EPE needs assessment and keep up the good planning undertaken for the Rome EPE	5/23/2018 9:48 PM
19	evaluating progress on SDGs taking into account each UN Agency country plans, the UNDAF and the national development plans	5/23/2018 5:33 PM
20	It would be interesting to have a session on systematic reviews/ synthesis methodologies during the next EPE, but in general I would like to see practical topics on management issues discussed as with 2018 EPE.	5/23/2018 5:15 PM

21	The convenor and moderators need to make sure that they share good experiences relevant to the topic	5/23/2018 4:39 PM
22	Perhaps more topics on methodologies and techniques with practical examples for UN agencies.	5/23/2018 4:26 PM
23	Na	5/23/2018 4:16 PM
24	replicate what was done this year!	5/23/2018 6:35 AM

Q32 Did you attend the AGM?

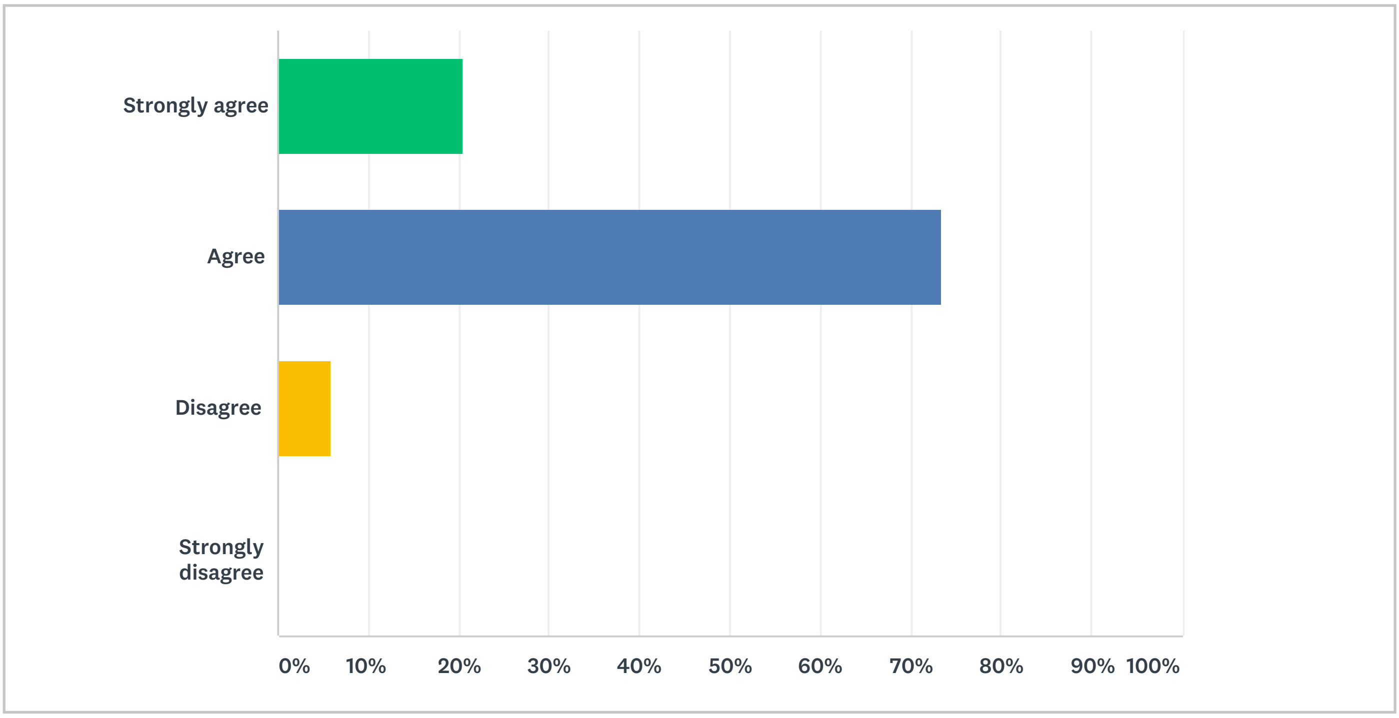
Answered: 63 Skipped: 10



ANSWER CHOICES	RESPONSES	
Yes	55.56%	35
No	44.44%	28
TOTAL		63

Q33 The time allocated to the AGM was adequate

Answered: 34 Skipped: 39

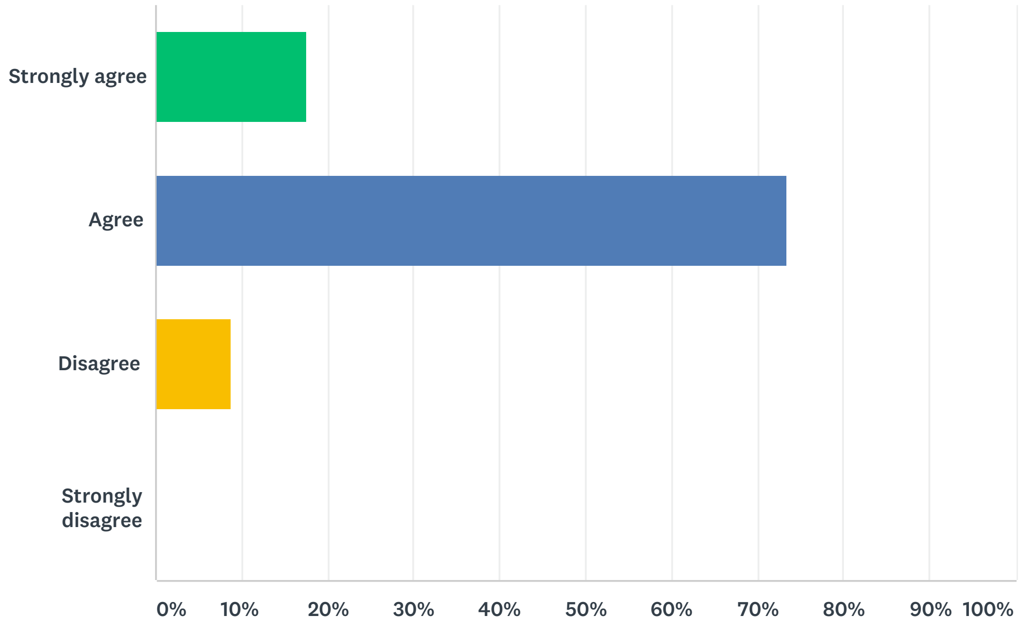


ANSWER CHOICES	RESPONSES	
Strongly agree	20.59%	7
Agree	73.53%	25
Disagree	5.88%	2
Strongly disagree	0.00%	0
TOTAL		34

#	COMMENTS	DATE
1	Time was adequate, but it was not results oriented	6/5/2018 9:52 AM
2	I did not participate in second day, so don't really know	6/5/2018 8:44 AM
3	Two days seems to be a good length.	5/26/2018 2:13 AM
4	Some discussions were rushed.	5/24/2018 10:18 PM

Q34 The AGM was useful to review progress and results of the work programme and decide on strategies and work areas for the following years

Answered: 34 Skipped: 39



ANSWER CHOICES	RESPONSES	
Strongly agree	17.65%	6
Agree	73.53%	25
Disagree	8.82%	3
Strongly disagree	0.00%	0
TOTAL		34

Q35 Do you have any comments on the AGM that could be shared or considered in future planning? (Duration, format, relevance, etc.)

Answered: 17 Skipped: 56

#	RESPONSES	DATE
1	The pre-meeting on the mid-term review of the UNEG strategy the day before was useful without which the discussions at the AGM would not have been as focused and productive.	6/6/2018 8:38 PM
2	Consider adopting a more informal approach, sharing the presentations ahead of the sessions and more accountability from each of the co-Chairs.	6/5/2018 5:56 PM
3	The AGM should be the culmination effort of a round year progress. Issues need to be discuss and negotiated before the AGM. The AGM should have a ratification role in the agenda and focus in one or two strategic issues to make more operational	6/5/2018 3:20 PM
4	The MTR was not used for decision making - the MTR was in fact side lined, a lost opportunity. While there were references to it, it should have been clear that all AGM members ought to fully read the recommendations beforehand and the jointly devise a response or path forward for UNEG, on the basis of these accepted or not accepted recs	6/5/2018 9:52 AM
5	this year, rather disjointed due to the MTR's influence on the agenda / discussions	6/5/2018 9:21 AM
6	Format did not allow for conclusive wrap up after each session	6/5/2018 8:44 AM
7	Possibly discuss long term strategy before mid-term work plan	6/4/2018 5:48 PM
8	It would have been good to engage the two consultants that had conducted the mid-term review to present their report and then engage the AGM in a discussion.	5/28/2018 2:51 PM
9	Some of the points were discussed in previous AGMs as well. Some presentations (eg partnerships) were not really good in terms of relevance and content.	5/26/2018 2:13 AM
10	though no closed sessions anymore, the sitting setting was still very "head-focused" - not sure what a good alternative would be though	5/25/2018 10:07 AM
11	No	5/24/2018 10:18 PM
12	no	5/23/2018 10:30 PM
13	Not really; perhaps better synchronization between the EPE and the AGM, so that the latter can feed better into the work discussed during the AGM	5/23/2018 9:49 PM
14	I believe less is more so UNEG AGM should focus more on the strategic direction, be more innovative in approaching challenges facing the UNEG Community, leverage existing partnerships to achieve its strategic goals. Agenda should be set to enable its members to have forward looking discussions and not focus too much on inward looking structural issues.	5/23/2018 4:49 PM
15	The MTR should have been the basis for discussion. The decision not to use the report was taken by a group that did not represent all UNEG heads	5/23/2018 4:45 PM
16	It would have been good to spend more time on discussing the results of the working groups even if this year I understand it was an exceptional year because of the mid-term review.	5/23/2018 4:17 PM
17	Needs to learn from the EPE to be more interactive and less structured.	5/23/2018 6:37 AM

Pibor

www.unevaluation.org • @un_evaluation • #UNEGweek