

Report of the United Nations Evaluation Group Evaluation Practice Exchange 2016 Seminar

Final

Dates: 25-26 April 2016

Location: World Intellectual Property Organization Headquarters, Geneva

Lead Agencies: ILO, UN Women, UNICEF, WFP, UNITAR, GEF, OCHA

This report, prepared by the EPE Management Group, encompasses the discussions and outcomes from the UNEG EPE 2016.

Acronyms and abbreviations

CALL	(Syria) Coordinated Accountability and Lessons Learned
EPE	Evaluation Practice Exchange Seminar
ESCWA	Economic and Social Commission for West Asia
FAO	Food and Agriculture Organization
GEF	Global Environment Facility
HEIG	Humanitarian Evaluation Interest Group
IAEG-SDGs	Inter-Agency Expert Group on Sustainable Development Goal Indicators
IFAD	International Fund for Agricultural Development
ILO	International Labour Organization
ISWE	Independent System-wide Evaluation
MDG	Millennium Development Goal
NECD	National Evaluation Capacity Development
OCHA	Office for the Coordination of Humanitarian Affairs
OECD-DAC	Organisation for Economic Co-operation and Development -- Development Assistance Committee
OHCHR	Office of the United Nations High Commissioner for Human Rights
SDG	Sustainable Development Goal
UN	United Nations
UN Women	United Nations Entity for Gender Equality and the Empowerment of Women
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNITAR	United Nations Institute for Training and Research
UNODC	United Nations Office on Drugs and Crime
UN-SWAP	UN System-wide Action Plan on Gender Equality and the Empowerment of Women
WFP	World Food Programme

Acronyms and abbreviations	2
Introduction	4
Stream 1. Evaluability of the Sustainable Development Goals (SDGs): Tackling evaluation of the SDGs—current approaches and lessons	4
Stream 2. No one left behind: The role of evaluation in ensuring equity, equality and non- discrimination	4
Stream 3. National evaluation capacity development for evaluation of SDGs and partnerships: Lessons and ways forward	4
Stream 4. Humanitarian agenda and SDGs.....	5
Stream 1. Evaluability of the SDGs: Tackling evaluation of the SDGs—current approaches and lessons	6
Session 1: Evaluability of the SDGs	6
Session 2: Evaluability of the SDGs—World Café	7
Session 3: Findings from a review of evaluability literature and past evaluation experience to inform UNEG’s contribution to the 2030 Agenda: Presentation and discussion	8
Stream 2. No one left behind: The role of evaluation in ensuring equity, equality and non- discrimination	11
Session 1: No one left behind: Equity and equality	11
Stream 3. National evaluation capacity development for evaluation of SDGs and partnerships: Lessons and way forward.....	13
Session 1: National evaluation capacity development for evaluation of SDGs and partnerships....	13
Session 2: World Café	14
Stream 4. Humanitarian agenda and the SDGs	15
Opening panel: Strategic issues around humanitarian evaluation and the SDGs	15
Session 1: Syria Coordinated Accountability and Lessons Learned (CALL) Initiative: Results from the evaluation synthesis and gap analysis	17
Session 2: Humanitarian stream: Ethics and humanitarian principles in humanitarian evaluation.	18
Session 3: Use of humanitarian evaluation: Key challenges and possible solutions	22
Session 4: Humanitarian evaluation methods	23
Session 5: Independent system-wide evaluation (ISWE) of UN operational activities for development: Methodological lessons learned from the pilot evaluations.....	26
Annex 1. List of participants: Registered participants, provided by UNEG Secretariat	30

Introduction

The United Nations Evaluation Group (UNEG) Evaluation Practice Exchange Seminar (EPE) 2016 was held at World Intellectual Property Organization Headquarters in Geneva from 25 to 26 April 2016. The meeting was organized by the UNEG EPE Committee¹ and opened by Francis Gurry, World Intellectual Property Organization Director and Colin Kirk, Director, Evaluation Office, United Nations Children’s Fund (UNICEF).

The 2016 EPE included the following four streams:

Stream 1. Evaluability of the Sustainable Development Goals (SDGs): Tackling evaluation of the SDGs—current approaches and lessons

This stream focused on sharing experiences in evaluability approaches and methodology. The panel covered a wide range of issues on evaluability ranging from: (a) challenges and lessons from the Millennium Development Goals (MDGs) and “Delivering as One” evaluation experiences; (b) collaborative approaches to tackle harmonized evaluation questions of cross-cutting themes; (c) ensuring evaluation is embedded in the design of interventions supporting SDGs; and (d) having evaluation offices “fit for purpose” for evaluating SDGs and mapping out potential data sources. Other issues on evaluability were suggested by interested agencies, including recent experiences in evaluability exercises.

Stream 2. No one left behind: The role of evaluation in ensuring equity, equality and non-discrimination

This stream focused on sharing experiences related to evaluation methodologies and approaches to ensure “No one is left behind”. It also focused on the role of evaluation in ensuring the SDGs are evaluated making certain social equity, equality and non-discrimination are included. The panel shared experiences from United Nations (UN) evaluation offices on different approaches to evaluating the SDGs that address social equity issues, gender equality and non-discrimination, as well as methodologies on addressing social equity issues in evaluation. Breakthroughs in social equity and gender equality approaches to evaluation were also encouraged.

Stream 3. National evaluation capacity development for evaluation of SDGs and partnerships: Lessons and ways forward

This stream focused on sharing experiences and lessons on national evaluation capacity development (NECD) with a view to inform future work on NECD for the achievement of the SDGs. The session discussed important questions including: What are the challenges for evaluating the SDGs at the national level? What are the capacities needed? What is the role of UNEG and of UN agencies in NECD in the new context? How can NECD contribute to the achievement of the SDGs with equity? How can UN agencies work together towards more effective NECD support? What have

¹ The UNEG EPE Committee was co-convened by Ada Ocampo (UNICEF), Florencia Tateossian (UN Women), Naomi Asukai (ILO), Anne-Claire Luzot (WFP), Brook Boyer (UNITAR), Geeta Batra (GEF), and Victoria Saiz-Omenaca (OCHA).

we learned from previous experiences on NECD? The goal of the stream was to conclude with action points for the future. Other aspects also addressed included: (a) EvalSDGs² as a renewed approach for advocating evaluation in the SDG processes; (b) supporting national evaluation systems; (c) documenting good practices; (d) promoting South-South and triangular cooperation; and (e) engaging non-traditional actors through innovative approaches. The session was initiated with a panel discussion where global experiences were presented and questions related to NECD in the frame of the “No one left behind” Agenda were discussed. The second part of the session followed the approach of a World Café where UN agencies shared experiences and lessons on NECD. In the latter part of the session, presenters and participants were engaged in an interactive wrap-up.

Stream 4. Humanitarian agenda and SDGs

Building on UNEG EPE 2014, this stream was intended to share experiences on evaluation of humanitarian action, discussing in particular: (a) the humanitarian-development interface and the challenges of assessing SDG progress in humanitarian contexts; and (b) evaluating humanitarian principles. The format of the stream was interactive and other issues were suggested by interested agencies.

² EVALSDGs is a network formed to add value and learning to SDGs, as well as support processes to integrate evaluation into national and global SDG review systems.

Stream 1. Evaluability of the SDGs: Tackling evaluation of the SDGs—current approaches and lessons

Session 1: Evaluability of the SDGs

Presenters:

- 1) Carlos Rafael de Medina Suarez, International Labour Organization (ILO) Director of Statistics and member of Inter-Agency Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs)
- 2) Colin Kirk, Director Evaluation Office, UNICEF

Presentation: [SDG Global Indicator Framework](#)

An agreement on a Global Indicator Framework for monitoring progress towards achieving the 17 SDGs and 169 targets was reached in March 2016. Although owned by nation states, the framework was proposed by the Inter-Agency Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs) and was adopted by the UN Economic and Social Council in June 2016. It is expected to be adopted by the General Assembly in September 2016. IAEG-SDGs is composed by Member States and includes regional, inter-agency and civil society as observers. IAEG-SDGs will monitor progress until 2030, drive the annual report, and review capacity development activities in statistical areas. It was noted that the indicators developed do not necessarily cover all aspects of goals and targets; data for several targets remain unavailable; and targets are not always quantified. It was also noted that there needed to be disaggregation by: income, gender, age, race, ethnicity, migratory status, disability, geographic location and national specific characteristics. More indicators can be added in the future as relevant to a specific context.

Indicators

Indicators are classified into three tiers:

- Tier I: An established methodology exists and data are already widely available
- Tier II: A methodology has been established but for which data are not easily available for all regions
- Tier III: An internationally agreed methodology has not yet been developed or it is not widely accepted or discussed and scarce information is available

Monitoring and reporting

Monitoring and reporting needs to occur at the national, regional and global level. Monitoring and reporting is voluntary, but there is a peer review system. A high-level political forum will be informed by an annual progress report on the SDGs, based on the global indicator framework. It will oversee follow-up and review and will conduct national reviews starting with 22 in July 2016, which are voluntary and driven by countries

Baseline approaches remain to be defined: Are they multi-indicator or indicator-specific ones? Are they national or regional? The baseline year for global tracking of MDG indicators was established in 1990, but there is no specific year set up for baseline of the SDGs—what should it be? The implications for the UN Development Assistance Frameworks should also be noted.

Implications

In order to fill the huge data gaps, capacity building in member states and constituents is important. However, resource constraints remain an issue, especially when considering global partnerships, data revolution and national budgets.

The Secretary-General's Report³ reflects an ambitious agenda behind five words (People, Planet, Prosperity, Participation, Partnership) and indicates the need for an increase in the number of actors and private sector involvement. To truly achieve the agenda of "No one left behind" means addressing the needs of one billion people amidst increased conflicts and climate change (rather than the growth environment of the MDGs). Business as usual will not suffice. This all has implications on the evaluability of the SDGs: By 2030, will we know what was achieved and how? We will need to know what changes have occurred, how they have occurred, and what their consequences are. Hence, there is a need for data and to come up with an imaginative evaluation approach.

Agency-specific evaluations might be comforting, but they are not enough to truly report on progress toward the SDGs. Relying on agency-specific evaluations could potentially replicate the programmer/operations silos often seen in evaluations as hindering progress. The SDGs are indivisible and interlocked; instead of a linear pathway of interventions, one might need to apply complexity theories to evaluation. How do we address the natural tendency of intersectional approaches to evaluations? We need new agile methodologies with adaptive step-by-step approaches. We also need to broaden our toolkits. Robert Chambers' advice is for eclectic pluralist approaches. One also needs to think about scale, given the massive development agenda ahead. Who performs evaluation and creates space for evaluation? We need to broaden the skills and roles of potential evaluators (e.g., evaluative activists such as the Global Evaluation Agenda of UNEG as well as national and local level evaluation champions). If the 2030 Agenda is to be properly evaluated, evaluative space needs to be created—and this requires partnerships, advocacy and communication.

Session 2: Evaluability of the SDGs—World Café

Presenters and presentations:

- 1) Sonya Meyerson-Knox, UN Economic and Social Commission for West Asia (ESCWA), *In-house meta-analysis of ESCWA evaluation*
- 2) Carlos Tarazona, Food and Agriculture Organization (FAO), *FAO evaluating UN agencies contributions to the SDGs: The FAO experience*
- 3) Johanna Pennarz, International Fund for Agricultural Development (IFAD), *IFAD/FAO/World Food Programme (WFP) evaluability of the SDGs: The main outcomes of an international workshop on the evaluability of SDG-2 held in Rome in November 2015 (video)*
- 4) Peter Wichmand, [ILO evaluability of SDGs through a decent work lens: ILO's work on evaluability of SDGs](#)
- 5) Barbara Torggler, UN Educational, Scientific and Cultural Organization (UNESCO), *UNESCO reflections about cultural context and the contribution of culture to the SDGs*
- 6) Anna Vighh, Global Environment Facility (GEF), *GEF evaluability of the SDGs related to saving the planet (Note)*
- 7) Indran Naidoo, UN Development Programme (UNDP), [UNDP evaluability of Goal 16 and challenges](#)

Representatives from ILO, FAO, IFAD, WFP, GEF, ESCWA, UNESCO and UNDP met in a "World Café" format to present their specific challenges in linking *inter alia* agency outcomes with the SDG goals.

³ A/RES/70/1 - Transforming our world: The 2030 Agenda for Sustainable Development, available online at: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

This alignment often requires rethinking the way we are working and collaborating. Important issues to address include:

- How to ensure proper monitoring takes place to inform evaluations
- New modalities and creative approaches to country-led evaluations
- Working more closely with statistics colleagues in each organization
- The need for clear indicators at the country level
- Practical considerations in dealing with complexities
- Challenges in mainstreaming the integration of evaluation in how an agency works on the SDGs
- Building blocks for the evaluation systems at various levels require involving many partners and a broadened scope
- Keeping the focus of evaluation as a public good and transformative
- The role of UNEG in all of the processes

Session 3: Findings from a review of evaluability literature and past evaluation experience to inform UNEG's contribution to the 2030 Agenda: Presentation and discussion

Moderator: Helen Wedgwood, Director WFP Office of Evaluation on behalf of the SO3 SDG Working Group

Presenter: Tullia Aiazzi, Independent Consultant

Presentation: [*Evaluation in the SDG era: Lessons, challenges and opportunities for UNEG*](#)

The aim of this presentation was to inform EPE participants about the key issues in the SDGs that are relevant to the UN evaluation system. The focus of the presentation was on the findings of a desk review commissioned by the SO3 SDG Working Group. The purpose of the desk review included the following: (a) an analysis of the evaluability challenges, opportunities and issues to consider in the 2030 Agenda based on a review of the evaluability literature, the 2030 Agenda documentation, stakeholder analysis, and ongoing indicator development work; and (b) a review of recent evaluation experiences and reports pertaining to MDG themes and/or selected country-level evaluations related to the MDGs. The desk review intended to derive lessons to support UNEG's evaluative efforts in respect to the 2030 Agenda and provide advice on what UNEG and its members should consider when: (a) developing future evaluation strategies, plans, approaches and methods; and (b) considering contributions to a shared global SDG evaluation agenda, including the potentials and risks of harmonized approaches.

Methodology

The methodology used for this desk review consisted of: (a) extensive document search and analysis, including all available evaluation policies of UNEG members; (b) interviews within UNEG and other stakeholders; (c) participation in the high-level event: *Leaving no one behind—Evaluating the Sustainable Development Goals with an equity-focused and gender-responsive lens* held in New York in March 2016; and (d) iterations with the Task Team, Working Group and peer reviewers.

The Inception Report identified the analytical framework for the review. The Organisation for Economic Co-operation and Development's (OECD) Development Assistance Committee (DAC)

definition was proposed (“*Evaluability means the extent to which an activity or project can be evaluated in a reliable and credible fashion*”) accompanied by the concept that an evaluability assessment should: (a) assess design, data and demand for an evaluation; (b) be in line with UNEG Norms 7.1⁴ and 7.2⁵; (c) take into account UNEG members’ experience with evaluability assessments; and (d) define the terminology, principles, roles and responsibilities.

The review was conducted in early 2016, at the beginning of the implementation phase of the 2030 Agenda, when all stakeholders were just starting to deal with the challenges posed by the SDGs. Therefore, no references or examples could be found of concrete initiatives contributing to the SDGs. However, the review did draw upon lessons from past evaluations including: UNEG members did not evaluate the MDGs; there is no complete single repository of UNEG evaluations; common methodological features and related lessons; and lessons for UN organizations and member states.

NECD: Responsibilities, demand and supply

UNEG should take the lead role in commissioning a world-wide mapping of national evaluation systems, building on the existing material, with a view to assessing the need for NECD and potential pathways of collaboration on country-led reviews. If a Member State will now be responsible for evaluation at the country level, this doesn’t mean that the UN system won’t do any more evaluations at the country level. Therefore, it is important to identify who will take the responsibility of the evaluations’ design at the country level. In this regard, there is a strong need for sharing and agreeing on common evaluation frameworks.

On the demand side, a huge gap in information availability has been identified. On the supply side, it has been observed that the supply is increasing. This would inevitably mean that UNEG may wish to collaborate with the national evaluation systems to carry out joint evaluations.

Human rights, gender equality and sustainable development are considered both goals and cross-cutting issues in the 2030 Agenda and will have to be integrated as such in evaluations. One crucial cross-cutting issue is still left behind: sustainable development. The review identified a lack of tools and methods to assess sustainable development.

SDG principles and evaluating the SDGs

The following SDG principles were identified:

- **Interlinkages:** SDGs are integrated and indivisible. The explicit recognition of the huge complexity of the SDG agenda is the new normal. There is a strong need to identify new and different ways to analyse inter-linkages among the 17 goals.
- **Universality:** Every Member State has committed to achieving the SDGs and implementing the 2030 Agenda. UN delivery is increasingly focused on normative work, global public good and capacity development.
- **Partnerships:** This is a key goal and implementation tool for the 2030 Agenda and evaluation. We must draw on UN evaluation system experience through joint evaluations at national,

⁴ **UNEG Norm 7.1:** During the planning stage of an undertaking, evaluation functions can contribute to the process by improving the ability to evaluate the undertaking and by building an evaluation approach into the plan. To safeguard independence this should be performed in an advisory capacity only.

⁵ **UNEG Norm 7.2:** Before undertaking a major evaluation requiring a significant investment of resources, it may be useful to conduct an evaluability exercise. This would consist of verifying if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators, assessable reliable information sources and no major factor hindering an impartial evaluation process.

regional, thematic and global levels; inter-agency humanitarian evaluations; and independent system-wide evaluations. EvalPartners is also a crucial achievement for the UN evaluation system. Finally, new actors need to be involved, including multi-stakeholder partnerships and major groups.

An implicit question in the Terms of Reference was whether or not the SDGs could be evaluated. The review concluded that the SDGs can and should be evaluated, despite the complexities and challenges inherent in the task. The discussion should therefore focus on “how the SDGs can be evaluated” and what the evaluation community can do to ensure that the evaluations in the 2030 Agenda context will be credible, reliable and useful. Influencing elements are: (a) design, data and demand; (b) integration of principles into evaluation mandates, design and practice; (c) evaluations that “talk to each other”; and (d) the role of UNEG and its membership.

Stream 2. No one left behind: The role of evaluation in ensuring equity, equality and non-discrimination

Session 1: No one left behind: Equity and equality

Facilitators: Florencia Tateossian, UN Women and Ada Ocampo, UNICEF

Opening speaker: Maria Bustelo, Dean Delegate for Equality, Universidad Complutense de Madrid

Presenters and presentations:

- 1) Oanh Nguyen, UNDP, [*UNDP: UNDP's contribution to gender equality and women's empowerment \(2008-2013\): Analytical frameworks*](#)
- 2) Urs Nagel, UNICEF, [*Impact evaluations of the UNICEF-IKEA Foundation programme on improving adolescents lives in Afghanistan, India and Pakistan: Integrating an equity and gender equality focus ex ante in programme and evaluation design*](#)
- 3) Miguel Torralba, IFAD, [*Three examples of how IFAD- IOE \[independent evaluation office\] has contributed to the "No One Left Behind" commitment through evaluations*](#)
- 4) *Corporate-level evaluation on gender equality and women's empowerment in IFAD in 2010*
- 5) *Evaluation synthesis: IFAD's engagement with indigenous peoples*
- 6) [*Evaluation synthesis on FAO and IFAD's engagement on pastoral development*](#)
- 7) Sabas Monroy, Office of the United Nations High Commissioner for Human Rights (OHCHR), [*Integrating human rights and gender equality in evaluations in OHCHR*](#)
- 8) Sabrina Evangelista, UN Women, [*Results from the SWAP analysis*](#)

This stream focused on sharing experiences related to evaluation methodologies and approaches to “No one left behind”—ensuring social equity, equality and non-discrimination.

Maria Bustelo from the Universidad Complutense de Madrid introduced the session encouraging the audience to think in transformative terms and to examine not only the differences between men and women, but also the underlying norms and structures that perpetuate inequality. She touched on the universality of the SDGs and their intersectionality in all areas as an important element to take into consideration when considering gender equality. She also stressed the importance of not waiting for programmes to integrate gender equality in order to evaluate from a gender perspective.

In the form of a “fishbowl” conversation, representatives from different agencies discussed findings from case studies and evaluations. Specifically:

- **FAO** presented two case studies from Myanmar and Guatemala that focused on indigenous groups. They found that even though project documents include provisions for examining indigenous groups, implementation doesn't always work this way. The takeaway? Evaluation teams should have good knowledge of local contexts and the right expertise in order to use an appropriate lens to evaluate. Furthermore, the evaluation process should be an empowering process.

- **UNICEF** discussed three impact evaluations and made the point that designing evaluations *ex ante* helps produce better evidence and also gives an opportunity to build in design features to address gender and inequality.
- **IFAD** presented three evaluations related to gender and indigenous groups. They stressed the need to develop comprehensive policies so that there are no gaps in coverage across organizations pertaining to gender and inequality.
- **UNDP** presented a rating scale for evaluation results to assess gender responsiveness with the use of a typology.
- **UNHCR** presented on the need to create good gender policy, with practice tips on including criteria, objectives and evaluation questions within the Terms of Reference.
- **UN Women** presented on the UN System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) assessment and stressed that, overall, UN entities are not doing well in terms of meeting the goals originally set for 2017. UN Women suggested external assessment to assess progress, as data of internal assessment shows over-reporting.

Stream 3. National evaluation capacity development for evaluation of SDGs and partnerships: Lessons and way forward

Session 1: National evaluation capacity development for evaluation of SDGs and partnerships

Moderator: Indran Naidoo, UNDP

Presenters and presentations:

- 1) Florencia Tateossian, UN Women, *Results from the New York event on evaluating SDGs with an equity and gender equality lens*
- 2) Ada Ocampo, UNICEF, *Lessons from NECD support experiences globally: How can experience inform capacity development in the context of the “No one left behind” Agenda*
- 3) Johanna Pennarz, IFAD, *IFAD’s approach to NEC [national evaluation capacity]: Experiences from NEC support in Ethiopia and China*
- 4) Roberto LaRovere, UNDP, *UNDP commitments from NEC conference: Establishing a country level M&E [monitoring and evaluation] for SDGs*

NECD has gained important momentum. The new 2030 Agenda for Sustainable Development, the General Assembly Resolution on National Evaluation Capacity Development [Resolution 69/237](#), and the Global Evaluation Agenda 2016-2020 open new opportunities as well as new challenges for NECD support. For the first time in history, there are global frameworks to guide NECD strategies and initiatives. In line with this new context, demand for NECD has significantly increased. Governments, parliamentarians, civil society organizations as well as other actors are demanding assistance for the creation or reinforcement of national evaluation frameworks and systems in line with the “No one left behind” Agenda.

Key points raised in the discussion related to the increased inter-connectivity of issues that the SDGs will bring and the related complexity with implications on capacity development needs. New approaches will be needed for evaluating complex initiatives. It was stressed that SDG evaluations will have to be country driven and the role of the United Nations should be to provide support. This also raised the issue of demand for evaluation at the country level: Is it really seen as essential to good governance with accountability and learning elements, or is it a burden? Uruguay was cited as an example where there was strong political support, however the process was still complicated. The response of the United Nations and hence the UNEG should take this into account and work in a coordinated manner with multiple partners (from traditional to less traditional) without creating competition. This also led to discussions on mapping of country needs, coordinated responses and measurement of UN efforts to strengthen national evaluation capacity.

Session 2: World Café

Presenters and presentations:

- 1) Adan Ruiz, United Nations Office on Drugs and Crime (UNODC), *“Boutique approach”: An experience of UNODC to NEC [national evaluation capacity] by blending evaluation and substantive approaches to topics such as terrorism and organized crime*
- 2) Craig Russon, ILO, [*ILO’s experience in the capacity development of the tripartite \(workers, employers and government\) constituent partners*](#)
- 3) Brook Boyer and Elena Proden, United Nations Institute for Training and Research (UNITAR), *Professional development: Forward looking opportunities*
- 4) Inoussa Kabore, UNICEF, *Sustainable capacity strengthening approach to VOPEs [voluntary organizations for professional evaluation]: Management, partnerships, fundraising, governance*

Representatives from different UN agencies met in a “World Café” format to discuss examples, challenges and lessons learned in developing national evaluation capacity. Specifically:

- **UNODC** used an example from Morocco (boutique evaluation approach) to show that national evaluation capacity can be empowering as a tool for change.
- **UNITAR** spoke about the demand side for national evaluation capacity. They noted that a lot is already happening in the field, but gaps still exist and there is also a need to avoid overlap.
- **UNICEF** provided an example of how capacity development must also address systemic issues not just technical evaluation training.
- **ILO** did a demonstration of a game illustrating that people learn differently and that methods need to be adjusted to the audience.

Stream 4. Humanitarian agenda and the SDGs

The sobering reality for the most vulnerable women, children and men is a world affected by increasing and ever more complex crises, and the principles of humanitarian action are increasingly at risk. This is a critical part of the backdrop to the 2015 SDGs and the 2030 Agenda. Humanitarian action faces increased challenges and complexity that is stretching the global humanitarian assistance community far beyond the traditional immediate emergency response. The world needs greater effectiveness, innovation, coordination, connectedness and coherence of humanitarian response as never before. Evaluation of humanitarian action has a vital contribution to make to achieving the 2030 Agenda.

Opening panel: Strategic issues around humanitarian evaluation and the SDGs

Moderator: Helen Wedgwood, Director of Evaluation, WFP

Panelists:

- 1) Colin Kirk, Director of Evaluation, UNICEF
- 2) Ewen MacLeod, Head Policy Development and Evaluation Service, UNHCR
- 3) Victoria Saiz-Omenaca, Evaluation Office for the Coordination of Humanitarian Affairs (OCHA)

Presentations:

- 1) *The SDGs and implications for humanitarian action with a specific focus on country-led humanitarian and development process and the implications for humanitarian assistance*
- 2) *New expectations about evidence: Secretary-General Report for the World Humanitarian Summit*
- 3) *Preliminary key lessons learned from major inter-agency humanitarian evaluations ahead of the World Humanitarian Summit (May 2016)*

Background

During the EPE Annual General Meeting held in 2015, three agencies expressed interest in reflecting on humanitarian evaluation issues, and the Annual General Meeting confirmed the launch of the Humanitarian Evaluation Interest Group (HEIG) as part of the UNEG workplan for 2015-2016. This represented a significant milestone in the way in which UNEG members see evaluation in different contexts. The HEIG now includes 10 agencies that work together on various issues such as the evaluation practice around humanitarian principles. The role of evaluators to increase effectiveness and efficiency of humanitarian action in view of the 2030 Agenda is increasingly important and more and more recognized.

Humanitarian assistance in the context of the humanitarian agenda

The globe is increasingly interconnected, and current conflicts often cross national borders. Humanitarian crises in the world have shown that if we want to achieve the SDGs, then we need to be able, through humanitarian assistance, to address these protracted crises, conflicts and natural disasters. The SDGs are not only about “development” but also about what is “socially sustainable”.

While we tend to see the MDGs in isolation and as separate goals, the SDGs are interrelated and interlocking goals. This poses real problems to evaluators: How do we address complexity? How do

we address long-term processes? Interconnection of issues requires a coherent and integrated response, and therefore we need to change our mind sets to deal with this new reality.

The SDG Agenda, unlike the MDG Agenda, is very much seen as a country-owned agenda (incorporation of the approaches to SDGs into national parliament plans and disaster preparedness plans). The gap between humanitarian action and development action (inconsistency of interventions) should be bridged at the country level and through a higher degree of resilience, coherence and connectedness of interventions.

Expectations on evidence or about evidence are critical challenges for the evaluation community, and this is the right moment to frame this issue in the perspective of other UN initiatives. In the Secretary-General Report “One humanity shared responsibility for the World Humanitarian Summit”, the need for evidence-based investments is scattered throughout the document. The report links humanitarian action to achieving the SDGs and identifies five core responsibilities:

- 1) Political leadership to prevent and end conflicts
- 2) Uphold the norms that safeguard humanity
- 3) Leave no one behind
- 4) Change people’s lives—from delivering aid to ending need
- 5) Invest in humanity

The report has an Annex “Agenda for humanity” that sets 24 goals, 40 immediate objectives and 139 activities to support the objectives and goals. There are 14 different references to information, data and monitoring, mainly in core responsibility Number 3, “Leave no one behind”. In this report, we should look at issues that are already on our humanitarian evaluation agenda: centrality of protection, accountability to affected populations, etc. Interestingly, there is no reference to any “cluster” approach. There are three key areas (inclusion, collective outcomes and financial arrangements) that are of interest to evaluation in particular. Considerable space has also been given to the creation of collective platforms. Several new areas will require attention: data collection and sharing, the formulation of collective outcomes, multi-year plans and benchmarks. The evaluation community is therefore going to face an increasingly challenging environment and a challenging agenda set out (including both quantitative and qualitative targets).

Inter-agency humanitarian evaluations

These are an assessment of results achieved in response to an emergency. They are mandatory for TA L3. Evaluations are conducted 12 to 15 months after the emergency has been declared. So far, three evaluations have been conducted: Philippines, South Sudan, and the Central African Republic, plus the Syria Coordinated Accountability and Lessons Learned (CALL), which is a learning initiative that presents findings from 20 different evaluations on the Syria CALL. There will be a synthesis of learning of these three inter-agency humanitarian evaluations plus the Syria CALL.

These evaluations show different degrees in the level of achievements. Much progress has been made, but we still need to improve coordination. The declaration of the L3 has been found as a very useful tool to quickly raise international attention around an emergency, relevance of response (positive), coordination, leadership and application of the TA principles. A webinar was held in May on key findings of the Syria CALL.

Session 1: Syria Coordinated Accountability and Lessons Learned (CALL) Initiative: Results from the evaluation synthesis and gap analysis

Moderator: Helen Wedgwood, WFP, in coordination with UNICEF, OCHA, UNHCR and ALNAP

Presenters:

- 1) James Darcy, Senior Evaluator
- 2) Julia Betts, Senior Evaluator

Presentation: [Syria CALL Evaluation Synthesis and Gap Analysis \(ESGA\): Findings and conclusions](#)

Purpose, scope and methodology

Undertaken as part of the Syria CALL Initiative, this synthesis was based on a review of 24 publicly available evaluations and evaluative studies concerning the international response to the Syria crisis, covering the period 2012 to 2015. The goal of the synthesis was to provide an overview of results of learning and accountability efforts undertaken thus far, giving an indication of the evolution of the humanitarian response, its challenges and achievements, and identifying emerging convergent themes and areas of findings as well as gaps of topics in evidence of the public domain.

The scope included five regional countries hosting refugees and Syrian displaced populations, but the bulk of the evidence reviewed concerned Lebanon, Jordan and Syrian refugees. It did not include the response to refugees in Europe. The reference material was varied and diverse and included either multiple single agency reports; country or programme specific reports; as well as thematic reviews, studies and evaluations.

Evaluators took a narrative approach, based on judgment about significance and analysis of evidence strength. The synthesis was not a meta-analysis but a compilation of key findings under seven thematic clusters based on identified themes. Evaluators were looking for patterns and recurrent or contrasting findings. It was left to the judgement of authors on the relevance and significance of findings.

Emerging themes and gaps were easily recognizable and validated, but treatment of themes (depth, sub-themes coverage, etc.) varied. Hence, assessment of the depth of evidence was difficult as the variability made it difficult to calibrate and quantify statements.

Findings

The evaluators identified more than 30 findings, falling under the following seven themes.

- Contextual
 - Restricted humanitarian space and access in Syria
 - Lack of effective humanitarian access has been critical in limiting the delivery of aid
 - Direct targeting of civilians and civil objects in Syria is disregard by all sides of basic International Humanitarian Law norms
- Strategy, planning, coordination and leadership
 - Lack of overarching strategy was reported as well as inefficiencies stemming from confusion over respective roles and/or overlap , with improvements noted
- Programme delivery, effectiveness, coverage and quality
 - Little is assessed in terms of preparedness, and mainly by process rather than actual utility (but an emerging interesting metric for it might be organizational flexibility)

- Effectiveness was assessed largely by delivery of outputs against operational objectives, but there was a lack of clarity of (shifting) targets and baseline
- Agencies were slow to foresee the scale of the crisis
- Major coverage gaps inside Syria exist
- There is little evaluation of the quality of the response against standards
- Protection, vulnerability, advocacy and humanitarian principles
 - UN protection efforts of Syria are not well joined up
 - There is an emphasis on normative statements, however less emphasis in terms of actionable commitments
 - Refugee protection was relatively successful, given the lack of strong international legal frameworks
 - There was weak evaluation against humanitarian principles
- Targeting, community engagement
 - There were limitations of accountability in Syria; more could be done
 - Social cohesion issues need to be better addressed
- Staffing, partnerships and operational efficiencies
 - Efficiency proved hard to evaluate given available data and lack of clear comparators
 - Partnerships were rarely analysed in terms of their effectiveness or appropriateness
 - There was system-wide organizational overstretch in terms of human resources and capacity
 - Assessment, monitoring and evaluation
 - There was a weakness of assessment inside of Syria (linking to accountability issues)
 - There was a lack of data about refugee populations living in host communities
- Evaluation gaps
 - Analysis of the response inside Syria
 - Risk analysis, including attention to “do no harm” principles, and to staff and partner security
 - Analysis of shortfall implications (unmet needs)
- Evidence gaps
 - Financial management
 - Quality of aid and compliance with standards
 - Remote assistance
 - Resilience and transition programming
 - Humanitarian principles: Application and compliance
 - Preparedness and organizational readiness
 - Advocacy and influencing strategies
 - Monitoring, reporting and information management
 - Granularity and information related to several specific groups including women, minorities, non-Syrian refugees, etc.

Session 2: Humanitarian stream: Ethics and humanitarian principles in humanitarian evaluation

Moderator: Anne-Claire Luzot, WFP, for the SO3 Humanitarian Evaluation Interest Group (HEIG)

Presenters:

- 1) Margie Buchanan-Smith, Independent Consultant
- 2) Hugo Slim, Head of Policy at the International Committee of the Red Cross

Purpose and methodology

The purpose of this desk review was to provide the HEIG with a better understanding on how the four core humanitarian principles of *humanity, impartiality, neutrality* and *independence* are evaluated through: a review of existing practice; and a reflection on the way forward and identification of approaches to ensure that these principles are included in the scope of evaluation of humanitarian action.

As a first step, an Inception Report was prepared including a data collection matrix, an analytical approach, and an outline of the report. The sample of 142 evaluations was screened using 10 key words: Humanity, Impartial(-ity), Neutral(-ity), Independen(-t, -ce), Dignity, Access, Space, Security, and Military.

The final report was built on five main data sources: (1) a review of general literature on the Humanitarian Principles since 2000; (2) analysis of humanitarian strategies, evaluation policies and guidelines of 10 different agencies; (3) a review of a sample of 142 evaluations covering 7 emergencies (Afghanistan, Democratic Republic of the Congo, Haiti, Somalia, Sudan, South Sudan and Syria); (4) analysis of a purposive sub-sample of 20 evaluations that made greater reference to humanitarian principles; and (5) interviews with 12 key stakeholders for reflections on humanitarian principles in evaluation of humanitarian action.

Findings from the evaluation sample

Overall

Humanitarian principles are not systematically assessed in evaluation of humanitarian action, either individually or as a group. The words “access” and “security” are the most frequently used (29 per cent of the total), but their link to humanitarian principles results is tenuous and implicit. Policies are not providing adequate direction to agencies, and there is a major gap in single and inter-agency evaluation guidance concerning evaluation of humanitarian principles.

Specifically

- The humanitarian principles are explicitly mentioned in about one-third of evaluations, but often an in-depth analysis is lacking
- “Impartiality” is the most frequently referenced principle, usually addressed under evaluation criterion of coverage
- Comprehensive evaluation of humanitarian principles combined is not taking place: most frequent combinations include “Independence and Neutrality” and “Independence and Impartiality”
- There was a concentration of terms in 20 out of 142 evaluations

Findings from the evaluation sub-sample⁶

Overall

⁶ The sub-sample is made up of the following: 12 reports from UN agencies, 5 from donors, 2 from International Federation of Red Cross and 1 from non-governmental organizations.

There is a lack of guidance on what constitutes good practice. Humanitarian principles are not systematically assessed in evaluation of humanitarian action.

Specifically

- There is no significant difference between different types of agencies in their treatment of humanitarian principles in evaluation of humanitarian action.
- Strategic and thematic evaluations are more likely to reference humanitarian principles than operational, RTE or impact evaluations.
- There are extremely limited examples of good practice (6 out of 142 evaluations).

Challenges to evaluating humanitarian principles

- There is no common understanding of the humanitarian principles in terms of concepts and implementation, which will make it challenging developing common guidance on how to evaluate them.
- Contradictions between the humanitarian principles as well as the existence of other potentially “conflicting” principles make evaluation of humanitarian principles even more challenging, especially where agencies are working to implement these different sets of principles simultaneously.
- Humanitarian agencies’ embarrassment with the increased politicization of humanitarian aid have produced the tendency for some agencies to undertake in-house evaluations not widely shared in the public domain in order to steer away from the challenges involved in evaluating their implementation of humanitarian principles. This can result in “behind closed doors” discussions.
- Methodological challenges: Standard evaluation methodologies must be adapted for evaluating performance against humanitarian principles. This would require introducing a more political lens into evaluations for overcoming the evaluation of humanitarian action tendency to be more technical.
- There is a lack of knowledge and expertise on humanitarian principles in evaluation of humanitarian action, both amongst evaluators and evaluation managers.
- There is also a lack of guidance on what constitutes “good practice”.

Lessons from the good practice examples

- Terms of References are not providing adequate guidance to evaluation teams, including in relation to evaluation questions. Therefore, Terms of References for evaluation of humanitarian action should include specific evaluation questions related to humanitarian principles, and include political context analysis (especially at the at inception phase) in relation to humanitarian principles.
- Methodology: Most are using standard methodologies, with one innovative example noted.⁷
- Two of the ECHO evaluations⁸ included a comprehensive analysis of implementation of humanitarian principles (thematic evaluations), bringing out tensions between principles.
- Detailed recommendations on humanitarian principles were not prevalent, although a minority of reports did include specific recommendations, mainly related to access and coverage.

⁷ The evaluation of the Disasters Emergency Committee (2001) response to the Gujarat Earthquake, which uses the Red Cross and Red Crescent Code of Conduct to evaluate the intervention.

⁸ ECHO commissioned two evaluations, which both include a specific focus on Humanitarian Principles: 1) Evaluation and Review of Humanitarian Access Strategies in DG-2012; and 2) Evaluation of the Implementation of the European Consensus on Humanitarian Aid-2014.

- There is a need for qualified expertise (Evaluation Manager and Evaluation Team) and commitment by organizations.
- There is also a need to take a different approach to evaluation of humanitarian action that is less mechanical, more imaginative and context specific.

Suggested follow-up for the HEIG

- Update inter-agency humanitarian evaluation guidance on large-scale system-wide emergencies to provide greater attention to evaluation of humanitarian principles, drawing on good practice examples when possible.
- Carry out a follow-up review of non-governmental organization evaluations assessing the extent to which they reflect humanitarian principles.
- Pilot evaluation of humanitarian principles (where there is a lesser degree of political conflict) drawing on available methodologies.
- Pilot use of a confidential Humanitarian Principles Annex in inter-agency and single agency evaluations, supported by a peer review to ensure that this is not used as a way of hiding unflattering results.
- Following ECHO and OCHA, commission single-agency evaluations focussing on humanitarian principles.
- Use existing Communities of Practice (UNEG, ALNAP, EvalPartners, etc.) to disseminate findings and move the debate and practice forward.
- Carry out meta-evaluations on a regular basis to assess if evaluation practice on humanitarian principles is improving.

Additional takeaways from the Head of Policy at the International Committee of the Red Cross

Humanitarian action is ethically grounded in humanitarian principles. We must look at the way we define what is “good practice”. Carrying out effective humanitarian action means to be in line with one principle. A set of principles can be in conflict with each other and bring some degree of incoherence. Humanitarian principles crash against hard politics and real problems. Balancing principles and interpreting what is independent and not independent is essential—and the job of evaluators is precisely to make this judgement.

Some principles can be in tension with each other—evaluations could look at the way operators make decisions and balance between those principles within a specific situation (What is neutral, independent, impartial enough? Did the operation focus on dignity and humanity?), given that operators make those judgement calls on a daily basis. Discussion around the humanitarian principles should include accountability to affected populations, community engagement, individual dignity, cultural sensitivity, coverage versus needs, etc., or to look at whether or not ethical conversations were held and if the right decisions were consequently taken. Another very important question to be answered is if the humanitarian action has improved or deteriorated the populations’ dignity.

We need to look at humanitarian principles and make judgements about how the humanitarian action was taking those principles into account in planning and implementation. Dignity is actually the key issue of accountability: Was the humanitarian operation dignified? Did it increase the dignity of people affected by the intervention?

Session 3: Use of humanitarian evaluation: Key challenges and possible solutions

Moderators:

- 1) Marta Bruno, FAO
- 2) Koorosh Raffii, UNICEF

Presenters:

- 1) Margie Buchanan-Smith, Independent Consultant
- 2) Sandra Aviles, Senior Advisor in Humanitarian Affairs, FAO Liaison Office
- 3) Kevin Savage, Humanitarian Research Director, World Vision International
- 4) Lori Bell, Evaluation, UNICEF
- 5) Krishna Belbase, Evaluation, UNICEF

Presentation: [Utilization of humanitarian evaluations](#)

Presenters addressed three important questions regarding humanitarian evaluation.

When in the evaluation process should utilization start?

Buchanan-Smith: As early as possible, already at design: e.g., in a learning-oriented evaluation, the Evaluation Team is invested in an important desk review, which is an initial investment to go into a country office with their findings to discuss the country team's views. This leads to increased ownership being built, but such processes need time and investments. Feedback from joint evaluations shows that individual agencies learn a lot from peer agencies—systemic issues should be better explored with joint evaluations approaches and/or synthesis and meta-reviews. We also need to find innovative ways of reporting and communicating evaluation practice. This is a great opportunity to draw down and find samples of positive practices and find the reason why certain circumstances are happening.

Savage: As a user of evaluation, involved in research-based programming, my focus is on organizational performance and learning. Use of evaluations is paramount to our decision-making processes, right at the beginning of the response. World Vision has no independent evaluation function; it is embedded within the programme section, and some of the most evaluative work is internal reviews at the country level. The process has often more importance for the country office; then the report's findings are used to disseminate learning at the central level and then replication in other responses at the decision-making phase.

Humanitarian evaluations are often country specific, but the findings and recommendations bring up issues that are systemic and recurrent. Can these issues bring traction for bigger change and how?

Savage: We found that our original evaluation cycle was too short and we have now adopted a three- to five-year evaluation cycle in order to have better evidence-based programming (it takes time to set a programme agenda). The evaluative agenda is also focusing on more specific topics that are the emerging recurrent themes from reviews and pairing up with research institutes for more systematic approaches to investigation.

Aviles: The HEIG and inter-agency humanitarian evaluation should feed into the Grand Bargain discussions (which is focused on efficiency and effectiveness, e.g. “accountability”). The Grand Bargain has been included in the Secretary-General’s Report for the World Humanitarian Summit as envisioned by the High-level Panel on Humanitarian Financing.⁹ A meeting in Washington, DC, last week looked at how to implement those commitments: The World Bank has been asked by one donor to come up with a mechanism to monitor the implementation. The World Bank presented its ideal process—an example of what the humanitarian community should adopt by including matrices from the analytical approach and evaluative processes. How is this group part of SDG and World Humanitarian Summit processes? Through the Grand Bargain discussions—but also through the new business model of joint outcomes—what is the rollout post the World Humanitarian Summit? (e.g., If multi-year programming were to become the rule, to what extent does the HEIG and inter-agency humanitarian evaluation [UN evaluation] need a champion to break through some of the current silos at highest level? For the roundtables underpinning each of the Grand Bargain core commitments, has the UNEG and inter-agency humanitarian evaluation the will to “sign up” for one of the core commitments? This could be a nice platform to bring it all together and to be taken up bilaterally by Assistant Secretary-Generals with Directors of Evaluation).

Who is the community of interest for humanitarian evaluation and what is the utilization balance between accountability, compliance, organizational performance and learning?

There is a need for systematic credible evaluations, but there is so much emphasis on the evaluation results and reporting phase (compared with field work and dissemination), that there must be other ways to enhance utilization (over an evaluation report or a video dissemination).

Session 4: Humanitarian evaluation methods

Moderators:

- 1) Marta Bruno, FAO
- 2) Victoria Saiz-Omenaca, OCHA

Presenters:

- 1) Koorosh Raffii, Krishna Belbase and Lori Bell, UNICEF
- 2) [Martha Bruno and Margie Buchanan-Smith, FAO](#)
- 3) [Claudia Martinez Mansell, Independent Consultant](#)
- 4) [Teresa Hanley, Independent Consultant](#)
- 5) Machiel Salomons, UNHCR

Inter-agency humanitarian evaluation of Typhoon Haiyan: Community consultation methodology

⁹ “Donors should provide multiyear funding, fewer conditions on donations and harmonise reporting requirements that can be very different from donor to donor. In return, humanitarian organisations reduce duplication and management costs, be more transparent about costs and commit to a “participation revolution” where beneficiaries are more listened to and included in basic operational decision making.”

In November 2013, an L3 emergency was declared. A four-person team conducted three weeks of field work in three types of communities and a half-day consultation to: (a) rebuild the community timeline; (b) discuss relevance, effectiveness, timeliness, transparency, qualitative and social differences; and (c) establish ratings as a group. The key principles underlying the approach included: (a) power of symbol, (b) power of position, (c) power of time, (d) power of space, and (e) power of numbers.

This consultation was complemented by other community activities to generate information and trends on community specific experiences. Several **constraints** were noted, including: a limited number of communities; a need for complementary acts to ensure coverage; community consultation fatigue; feedback to communities was limited to evaluation results; feedback from communities covered more than the humanitarian country team response and encompassed wider response; and the process required great community facilitators.

The following was **learned** from the process: there is the potential of working with large groups in a short time; there is value in scoring, but for trends and qualitative data only; factors that helped included data availability, community comfort with participatory processes, and great facilitators; we need to think through how to combine with accountability to affected populations processes through response and agency own evaluation; and what's the right balance—affected population input versus management issues?

Participatory impact assessment of FAO 2014 livelihood kits emergency distribution

In 2014, FAO reported to have supported the livelihoods of more than 400.000 households with different types of interventions, comprising livelihoods kits of staple crops seeds and tools, vegetable kits and tools and fishing kits / as well as distributing 800.000 kits of essential veterinary drugs and vaccinating 1.5 million animals.

Evaluators identified the following themes:

- Weak context analysis: What were the changes in livelihoods over time? There was an absence of monitoring data; what was the initial impact of the livelihood kits? What arrangements were made and what was the effectiveness of partnerships?
- Importance of team selection: FAO was using only nationals to ensure access, but it was difficult to get gender balance in teams; the inclusion of FAO monitoring and evaluation and gender officers; it was difficult to ensure coherence and similar skills sets across teams.
- Issues of ethics: There were issues in dealing with affected populations (ethical code of context for enumerators) and issues of coverage.
- Lessons on planning and methods: Logistics and flexibility—plan big; address seasonality and migration (people and livestock); triangulation and attribution are difficult; and language.

Balloon mapping: Evaluation methods

The following addressed Palestinian refugee camps in Lebanon. Available community maps and Google maps were not detailed enough, hence aerial photography mapping was conducted. The optimal uses for this approach is with concentrated populations and locations and fast occurring changes. The beneficiaries served as active actors in the mapping system (data collection process).

Transtec Lebanon and Jordan methodology

Engaging with refugees is sometimes difficult and sensitive. Transtec applied the OECD-DAC criteria but used many different tools, such as key informant interviews, surveys, focus group discussions, desk reviews, etc. The innovation was a household survey with persons of concerns framed on evaluation questions and risk assessment for each question. The enumerators (nationals with no Syria family ties) used PDAs to enter data with a daily central upload; ethical standards by HCR were addressed starting from the Terms of Reference.

Session 5: Independent system-wide evaluation (ISWE) of UN operational activities for development: Methodological lessons learned from the pilot evaluations ([PPT](#))

Presenters:

- 1) Members of the Independent System-wide Evaluation (ISWE) Evaluation Management Groups
- 2) Mike Reynolds, Evaluation Expert
- 3) Scott Green, Coordinator, ISWE Secretariat Joint Inspection Unit (Observer)

Presentations:

- 1) Scott Green, ISWE Secretariat Coordinator, *Background and objectives of the pilot ISWE mechanism*
- 2) Michael Reynolds, Independent Evaluation Consultant, *Results of the testing of the comprehensive evaluation modality*
- 3) Scott Green, ISWE Secretariat Coordinator, *Results of the testing of the synthesis evaluation modality*
- 4) Inga Sniukaite, UN Women and Krishna Belbase, UNICEF, *Brief perspectives from Evaluation Management Group members*
- 5) Inspector Sukai Prom-Jackson of the Joint Inspection Unit, *Moderated discussion and concluding remarks*

System-wide evaluation and the UN “fit for purpose” debate

The United Nations is seeking to position itself in a global context of massive and fast moving changes. In this context, it needs evidence-based information on its overall performance in programme countries and on how it should strategically position itself to add value in a context of multiple development partners. The ISWE initiative is set to provide evidence on whether or not the UN system is:

- Doing the right things in the right way
- Making a difference and adding value
- Enhancing the attainment of impact ensuring sustainability
- Could do things differently

The main objectives of the session were to:

- Provide a general briefing related to the ISWE policy
- Share methodological lessons related to the testing of two pilot ISWE evaluations
- Hear perspectives on ISWE from UNEG Evaluation Management Group members

The fragmentation challenge

- Currently, evaluations of operational activities by the various UN system organizations are segmented, dispersed, and cannot be easily aggregated or synthesized to support decision making about the UN system as a whole.
- Governing bodies as well as Member States of the UN system need more integrated and comprehensive knowledge about the effectiveness of UN operational activities for

development to improve coherence, support mutual accountability, and enhance effective decision-making at a system-wide level.

- ISWE supports the challenge of the SDGs by breaking down UN institutional barriers, encouraging holistic approaches and systems thinking, and as one of its fundamental principles, seeks to build national evaluation capacities.

ISWE definition and purpose

ISWEs apply a common framework to produce in-depth evaluations of specific policies, strategies, programmes, issues, efforts, areas or sectors in a single or in several countries, which are then synthesized into one study. The purpose of ISWE operational activities for development are to strengthen governing bodies in their role for oversight, accountability, decision-making, and direction setting of the UN system as a whole. ISWE supports system-wide coherence and contribution to development effectiveness and sustainability.

ISWE Objective 1: Functions, systems and mechanisms for management and for the conduct of evaluation

The ISWE policy recognizes the Joint Inspection Unit mandate for ISWE and seeks to strengthen the capacity for the conduct of ISWE of UN operational activities for development.

- It seeks to enhance partnership, making use of existing evaluation mandates, mechanisms and capacities of the Joint Inspection Unit and evaluation functions of UN system organizations
- It makes use of Evaluation Management Groups chaired by the Joint Inspection Unit and having membership of evaluators from UN system evaluation offices to enhance professional technical quality.
- It seeks to enhance and strengthen the strategic and substantive value, ownership and utility through establishment of Key Stakeholder Reference Groups.
- It makes use of external expert consultants to enhance the quality and efficiency of the evaluation.

ISWE Objective 2: Piloting of two modalities of system-wide evaluation

The second objective of the policy is to conduct two pilot evaluations addressing two of three identified evaluation modalities for system-wide evaluation. The two modalities piloted in 2015-2016 are: a comprehensive evaluation and a synthesis evaluation. The pilot evaluations evaluated whether or not: the UN system as a whole is responding to country needs and priorities and making progress towards internally agreed development goals; it does so in an appropriate manner that is efficient, coherent and ensures sustainability; and it is making a contribution to development results of programme countries and strengthening the capacity for impact. The pilot evaluations also provided information on what is required to carry out the two modalities of system-wide evaluation and the lessons for the future.

Results of the testing of the comprehensive evaluation modality

Pilot evaluation: Evaluation of the contribution of the UN development system to strengthening national capacities for data collection and analysis to support the achievement of the MDGs and other internationally agreed development goals

The pilot evaluation included approximately 30 UN system entities (specialized agencies, funds and programmes, regional commissions, UN secretariat) and a large number of national stakeholders (national statistical offices, line ministries, central banks, planning bodies) all producing statistics plus users in private sector and civil society. The evaluation covered the period from 2000 but realistically focused on the last five years. Approximately 120 UN country teams engaged in national statistical capacity development, thus the pilot project was not just about statistics but also about capacity development, UN county level coordination, partnerships and the political economy of use. The ISWE policy identified the following advantages and disadvantages in the approach:

Advantages: Generates new knowledge from purposefully selected cases, engages the partner countries extensively in the conduct of evaluations, and builds on and enhances national evaluation capacity.

Disadvantages: Costly, takes a long time, and requires an intensive coordination effort.

What happened during implementation—the policy was right! There was: not enough time (future ISWEs need to be undertaken over a period of years); uncertainty over money (budgets need to be secured up front); and administrative blockages (very strong dedicated evaluation management is required).

The approach was therefore dictated by these **realities**:

- Country studies as the core source of evidence are in line with the ISWE policy
- However the evaluation promoted little national evaluation capacity development
- Nor was there as extensive engagement with national partners in the countries as hoped
- Other sources of evidence also became very important, especially evaluations

The following **challenges to the design** were identified:

- How to ensure a full understanding of the context with such a broad scope
- How to ensure that you are able to collect enough evidence across the enormous scope
- Aggregating results: how to say something about the work of the UN system when results are inevitably mixed

Results of the testing of the synthesis evaluation modality

As per the policy, synthesis evaluation is one of three modalities for ISWE being tested under the pilot.

The following are seen to be the **main advantages** of the synthesis approach: they are desk studies that make use of existing evaluations from the UN system and other entities as well as relevant independent research; and they can be conducted relatively quickly and at low cost.

The main **disadvantage** is that they are dependent on existing materials, which may not be enough to adequately answer the questions and may require light complimentary data collection.

Key challenges

Resources were front loaded on this evaluation to allow its completion in 2015. As expected of a synthesis evaluation, there were many challenges that the Evaluation Management Group sought to manage in a proactive and pragmatic way.

- One major constraint was the **limited coverage of the underlying UNDAF evaluation set**: While it was easy to access a large number of UNDAF evaluation reports from the UNEG/UNDP website, evaluators could only find 27 formal evaluation reports (excluding MTRs). Given the UNDG requirement for all UNDAFs to undergo a formal evaluation, the team needed to confirm the extent of the compliance problem. Extra time was needed to conduct a survey of all UN country teams to be sure all available UNDAF evaluation reports were available. The survey resulted in an extra 9 UNDAF evaluation reports bringing the total to 36. A total of 88 countries, or 73 per cent of UN country teams with UNDAFs did not prepare the required evaluations.
- The **limited resources** available compounded the problem: There was an assumption that UNDAF evaluations would be the main source of underlying evidence. However, given the size of the UNDAF evaluation coverage gap and the limited resources at the team's disposal (only 28 per cent of what had been determined to be needed), it was not possible to fully make up for the gap through light complimentary data collection. The Evaluation Management Group decided to move ahead anyway in the spirit of experimentation. An initial evaluability study might have helped better define the approach that would be needed.
- **Aggregating results** at the impact level also proved challenging: Only 2 of the 23 validated UNDAF evaluations formally established any link between UNDAF activities and their impact on national poverty alleviation goals and strategies. It therefore proved challenging to aggregate any results at this high level of analysis. An approach at aggregating at the outcome level might have been more useful.

Lessons learned

Overall, the approach seems to have potential. However, before deciding on the feasibility of using a synthesis approach to ISWE in the future (which is assumed to be low cost), it would be useful to undertake a rapid evaluability assessment of whether or not the underlying evaluations and material can, in fact, provide the necessary coverage to answer key ISWE questions—a methodology primarily driven by synthesis using only light complimentary evidence. If not, then additional resources will be required with the synthesis approach becoming only one component of a broader methodology. The scope of the work and questions to be answered might alternatively be reduced or simplified.

Adopting a synthesis approach to aggregating a system-wide evaluation of UN performance at the global level is challenging, especially if the individual evaluations were never designed for this broader purpose. There will be gaps. In the future, a broader global UNDAF evaluation framework might help aggregate results better and ensure a more coherent set of evaluations.

Annex 1. List of participants: Registered participants, provided by UNEG Secretariat

No	First name	Family name	Organization	Email
1	Ruiz Villalba	Adan	UNODC	adan.ruiz-villalba@unodc.org
2	Tullia	Aiazzi	UNEG - consultant	tullia.aiazzi@gmail.com
3	Clairene	Alexander	UN	alexander6@un.org
4	Guido	Ambroso	UNHCR	ambroso@unhcr.org
5	Katherine	Aston	United Nations Volunteers	katherinemaston@gmail.com
6	Naomi	Asukai	ILO	asukai@ilo.org
7	Martin	Barugahare	UN-Habitat	martin.barugahare@unhabitat.org
8	Geeta	Batra	GEF	gbatra@thegef.org mkoukoui@thegef.org
9	Nivine	Bazzi	JIU - intern	nbaz8240@uni.sydney.edu.au
10	Krishna	Belbase	UNICEF	kbelbase@unicef.org
11	Lori	Bell	UNICEF	lbell@unicef.org
12	Luisa	Belli	FAO	luisa.belli@fao.org Yuenching.ho@fao.org
13	Elise	Benoit	UN World Food Programme	elise.benoit@wfp.org
14	Michaela	Berndl	United Nations Industrial Development Organization (UNIDO)	M.Berndl@unido.org
15	Simon	Bettighofer	WHO	bettighofers@who.int
16	Julia	Betts	UN-OCHA	julia@jbetts.net
17	Francesca	Bonino	UNHCR	bonito@unhcr.org
18	Brook	Broyer	UNITAR	brook.boyer@unitar.org
19	Marta	Bruno	FAO	marta.bruno@fao.org yuenching.ho@fao.org
20	Arnaud	Bubeck	WIPO	arnaud.bubeck@wipo.int
21	Margie	Buchanan-Smith	Independant evaluator	margiebsmith@aol.com
22	Sally	Burrows	WFP	sally.burrows@wfp.org
23	Maria	Bustelo	Complutense University of Madrid	mbustelo@ucm.es
24	Nan	Buzard	ICVA	nan.buzard@icvanetwork.org
25	Diana	Cartier	IOM	dcartier@iom.int
26	Gang (Daniel)	Chen	UNCTAD	daniel.chen@unctad.org
27	Andrea	Cook	UNFPA	acook@unfpa.org
28	Edgar	Dante	ESCAP	dante@un.org
29	Hicham	Daoudi	UNFPA	daoudi@unfpa.org
30	James	Darcy	UN-OCHA	jdarcywork@btinternet.com saiz-omenaca@un.org
31	Paola	Deda	UNECE	paola.dede@unece.org
32	Rafael	Diez de Medina	ILO	diez@ilo.org
33	Petru	Dumitriu	JIU	pdumitriu@unog.ch
34	Tuncay	Efendioglu	WIPO	tuncay.efendioglu@wipo.int
35	Zamira	Eshmambetova	UNECE	zamira.eshmambetova@unece.org
36	Sabrina	Evangelista	UN Women	sabrina.evangelista@unwomen.org
37	Erica	Fiekowsky	ILO	fiekowsky@iloguest.org

38	Christophe	Franzetti	IOM	cfranzetti@iom.int
39	Luisa	Frescura	UNAIDS	frescural@unaids.org
40	Susanne	Frueh	UNESCO	s.frueh@unesco.org
41	Ian	Fuller	International Criminal Court	ian.fuller@icc-cpi.int
42	Ricardo	Furman	ILO	furman@ilo.org
43	Andrew	Fyfe	UN Capital Development Fund	andrew.fyfe@uncdf.org
44	Oscar	Garcia	IFAD	o.gorcia@ifad.org l.danielsson@ifad.org
45	Geoffrey	Geurts	UNESCO	g.geurts@unesco.org
46	Josse	Gillijns	GEN	josse.gillijns@ifrc.org
47	Esther	Gomez	ILO	gomez@ilo.org
48	Scott	Green	JIU/ISWE Coordination Secretariat	Green10@un.org
49	Javier	Guarnizo	UNIDO	j.guarnizo@unido.org
50	Charlotte	Gunnarsson	UNODC	charlotte.gunnarsson@unodc.org
51	(Eddie) Yee Woo	Guo	OIOS	guoy@un.org
52	Teresa	Hanley	UN-OCHA	thanley7@gmail.com
53	Iris	Hauswirth	ITC UNCTAD/WTO	hauswirth@intracen.org
54	Stefan	Helck	JIU	shelck@unog.ch
55	Yuen Ching	Ho	FAO	yuenching.ho@fao.org
56	Masahiro	Igarashi	FAO	masahiro.igarashi@fao.org yuenching.ho@fao.org judita.jankovic@gmail.com jjankovic@icao.int
57	Judita	Jankovic	ICAO	
58	Miguel	Jimenez Pont	ITC UNCTAD/WTO	jimenez@intracen.org
59	Inoussa	Kabore	UNICEF	ikabore@unicef.org
60	Katharina	Kayser	UNODC	katharina.kayser@unodc.org
61	Nishta	Keeble	UNCTAD	nishta.keeble@unctad.org
62	Colin	Kirk	UNICEF	ckirk@unicef.org gdey@unicef.org
63	Jean-François	Klein	ILO	kleinj@ilo.org
64	Sonya	Knox	UN-ESCWA	knoxs@un.org
65	Katinka	Koke	UNITAR	katinka.koke@unitar.org
66	Roberto	La rovere	UNDP	roberto.la.rovere@undp.org
67	Jesus	Lara Alonso	JIU	jlara@unog.ch
68	Itziar	Larizgoitia	WHO	larizgoitia@who.int
69	Aurélie	Larmoyer	IAEA	a.larmoyer@iaea.org m.saqet@iaea.org
70	Paul	Lewe	JIU - intern	lewe.max@gmail.com
71	Emanuel	Lohninger	UNODC	emanuel.lohninger@unodc.org
72	Frederica	Lomiri	WFP	frederica.lomiri@wfp.org
73	Richard	Longhurst	GEN	r.longhurst@ids.ac.uk
74	Anne-Claire	Luzot	WFP	anneclaire.luzot@wfp.org
75	Ewen	Macleod	UNHCR	Macleod@unhcr.org
76	Stephanie	Martin	JIU - intern	smartin-intern@unog.ch
77	Claudia	Martinez Mansell	FAO	clauds@gmail.com Marta.Bruno@fao.org
78	Susanne	Mattsson	UNFPA	mattsson@unfpa.org
79	Martha	McGuire	Independant consultant	martha@cathexisconsulting.ca
80	Deborah	McWhinney	UN World Food Programme	deborah.mcwhinney@wfp.org

81	Maria Beatriz	Mello da cunha	ILO	cunha@ilo.org
82	Carsten	Meyer	IAEA	c.meyer@iaea.org m.saquet@iaea.org
83	Ljerka	Milos-King	JIU	lmilos-king@unog.ch
84	Sabas	Monroy	OHCHR	smonroy@ohchr.org
85	Helen	Morris	UNHCR	morris@unhcr.org
86	Delfina	Müller	JIU - intern	dmuller-intern@unog.ch
87	Janette	Murawski	ILO	murawski@ilo.org
88	Urs	Nagel	UNICEF	unagel@unicef.org
89	Indran	Naidoo	UNDP	indran.naidoo@undp.org
90	Björn	Neuhaus	GEN	bjoern.neuhaus@sfiwet-switzerland.ch
91	Oanh	Nguyen	UNDP	oanh.nguyen@undp.org
92	Ada	Ocampo	UNICEF	aocampo@unicef.org
93	Zenda	Ofir	GEN	evaluation@zendaofir.com
94	Alok Kumar	Ojha	WMO	aojha@wmo.int lcourital@wmo.int
95	Fumika	Ouchi	UNEG Secretariat	fumika.ouchi@undp.org
96	Salil	Panakadan	UNAIDS	panakadans@unaids.org
97	Oktavianto	Pasaribu	ILO	pasaribu@ilo.org
98	Rafael	Peels	ILO	peels@ilo.org
99	Elisabetta	Pegurri	UNAIDS	pegurrie@unaids.org
100	Johanna	Pennarz	IFAD	j.pennarz@ifad.org
101	Ramiro	Pizarro	ILO	pizarro@ilo.org
102	Jan	Pleser	JIU - intern	jpleser-intern@unog.ch
103	Marina	Ploutakhina	UNIDO	m.ploutakhina@unido.org
104	Elena	Proden	UNITAR	elena.proden@unitar.org
105	Sukai	Prom-Jackson	JIU - Inspector	promjackson@un.org
106	Koorosh	Raffii	UNICEF	kraffii@unicef.org
107	Rana	Raj	GEN	raj@theWolfGroup.org
108	Thierry	Rajaobelina	UNRWA	t.rajaobelina@unrwa.org
109	Patricia Hanitra	Ralijemisa	WMO	pralijemisa@wmo.int lcourital@wmo.int
110	Pascale	Reinke-Schreiber	UNODC	pascale.reinke-schreiber@unodc.org
111	Inmaculada	Roman	Evaluando	iroman@evaluando.es
112	Craig	Russon	ILO	russon@ilo.org
113	Victoria	Saiz-Omenaca	UNOCHA	saiz-omenaca@un.org
114	Machiel	Salomons	UNHCR	salomons@unhcr.org
115	Kevin	Savage	World Vision International	kevin_savage@wvi.org
116	Marianne	Schmitt	ITC UNCTAD/WTO	schmitt@intracen.org

11 7	Ekaterina	Sediakina- Riviere	UNESCO	e.sediakina@unesco.org
11 8	Marco	Segone	UN Women	marco.segone@unwomen.org
11 9	Anand	Sivasankara Kurup	WHO	sivasankarakurupa@who.int
12 0	Hugo	Slim	ICRC	hslim@icrc.org mpittetnazareno@icrc.org
12 1	Corey	Smith	ILO	smithco@iloguest.org
12 2	Inga	Sniukaite	UN Women	inga.sniukaite@unwomen.org
12 3	Michael	Spilsbury	UNEP	michael.spilsbury@unep.org
12 4	Robert	Stryk	UNRWA	r.stryk@unrwa.org
12 5	Isabel	Suarez	Un Women	isabel.suarez@unwomen.org
12 6	Georgiadis	Sylta	OHCHR	sgeorgiadis@ohchr.org
12 7	Carlos Antonio	Tarazona Cervantes	FAO	carlos.tarazona@fao.org
12 8	Florencia	Tateossian	UN Women	florencia.tateossian@unwomen.org
12 9	Kseniya	Temnenko	GEF	ktemnenko@thegef.org
13 0	Guy	Thijs	ILO	thijs@ilo.org
13 1	Saila	Toikka	UNEP	saila.toikka@unep.org
13 2	Rose	Tompson Coon	GEN	rose.thompson-coon@wsscc.org
13 3	Barbara	Torggler	UNESCO	b.torggler@unesco.org
13 4	Miguel	Torralba	IFAD	m.torralba@ifad.org
13 5	Inam	Ullah	UNECA	uinam@uneca.org
13 6	Geraldine	Vertallier	GEN	g.vertallier@worldvision.ch
13 7	Patricia	Vidal Hurtado	WIPO	patricia.vidalhurtado@wipo.int
13 8	Anna	Viggh	GEF	aviggh@thegef.org
13 9	Alexandra	Warner	ALNAP	a.warner@alnap.org
14 0	Helen	Wedgwood	WFP	helen.wedgwood@wfp.org luisiana.lopez@wfp.org
14 1	Peter	Wichmand	ILO	wichmand@ilo.org
14 2	Lukasz	Wieczerek	Organisation for the Prohibition of Chemical Weapon (OPCW)	lukasz.wieczerek@opcw.org
14 3	Janet	Wildish	UNEP	janet.wildish@unep.org
14 4	Jennifer	Worrell	OHCHR	jworrell@ohchr.org ciryama-mangin@ohchr.org
14 5	Yichen	Xu	JIU - intern	yxu-intern@unog.ch

14 6	Anida	Yupari	UNCTAD	anida.yupari@unctad.org ismahen.chouder@unctad.org
14 7	Amanuel	Zerihoun	JIU	azerihoun@unog.ch
14 8	Jin	Zhang	UNEG Secretariat	jin.zhang@undp.org
14 9	Jicheng	Zhang	ITC UNCTAD/WTO	zhang@intracen.org