

Annual
General
Meeting
2023



UNEG
United Nations Evaluation Group

Compendium of Evaluation Methods Reviewed - Volume II Exploring Methodological Appropriateness under COVID Conditions

December 2022

[DRAFT FOR COMMENT]

Dates: 24-26 January 2023

This document has been prepared by the UNEG Methods Use and Appropriateness Working Group for discussion under the Strategic Objective 2 session at the UNEG AGM 2023.

The WG invite comments on how to improve its structure, content and ultimately utility for the UNEG community.

See also:

- [Compendium of Evaluation Methods Reviewed \(Volume 1\)](#) (2020)

Table of Contents

INTRODUCTION	2
SETTING OUR PRIORITIES	4
METHODS DISCUSSED.....	4
1. REAL-TIME LEARNING APPROACHES.....	5
1.1. DEVELOPMENTAL EVALUATION.....	5
2. USE OF TECHNOLOGY AND SOCIAL MEDIA FOR EVALUATION DATA COLLECTION	9
2.1. GEOSPATIAL METHODS.....	9
2.2. SOCIAL LISTENING	11
3. STAKEHOLDER ENGAGEMENT IN EVALUATION	13
4A. ADVANCED ANALYTICAL TECHNIQUES THAT CAN BE APPLIED REMOTELY (PROCESS TRACING).....	14
4B. CASE STUDY AND CASE-BASED EVALUATIONS	17
CONCLUDING REMARKS	20
WORKPLAN OBJECTIVES 2023	20
SPEAKERS' SHORT BIOS	21

Foreword (zero draft)

Strengthening the professional capacity of UN evaluators to design and manage evaluations is a central objective of the United Nations Evaluation Group (UNEG) Strategy for 2020 – 2024 as UNEG aims to support the professionalization and capacities of its membership and safeguard and bolster the quality of evaluation practice in the UN system.

This in turn will help ensure that evaluations within the UN system can produce knowledge and evidence that can be used to inform relevant, effective, sustainable UN support to Member States to achieve the Sustainable Development Goals.

*The continuous effort to leverage evidence-based decision making at all levels of the delivery chain, whether humanitarian, development, or peace – remains a primary focus of UNEG. This can only be achieved through the application of credible evaluation methods. It is in this vein that I welcome the efforts of the UNEG Working Group on Methods. Volume II of the compendium of methods looks at four themes through the lens of specific evaluation methods: (i) **Methods and approaches that facilitate real-time learning** (developmental evaluation), (ii) **The use of technology and social media for [remote] data collection** (geospatial methods and social listening), (iii) **How best to support full participation of partners and beneficiaries, bearing in mind the centrality of the Leaving No One Behind objectives to the SDGs and the challenges posed by the pandemic to reach out the most vulnerable group** (stakeholder engagement), and (iv) **Methods and approaches to assess the causal relationship between what UN agencies deliver and the results that they claim at different levels of their impact pathways** (process tracing and case study and case-based approaches in evaluation).*

This document takes stock of how these methods have been applied in concrete UN evaluation contexts during the COVID-19 crisis, and is a useful resource for evaluators seeking to expand the range and diversity of methods and approaches in their evaluation work.

I would like to thank in particular the coordinators of the group: Tina-Tordjman-Nebe (UNDP), Henri van den Inert (UNHCR), and Francesca Carini (UNHCR); for carrying forward the work of the group in 2021 and 2022, and I look forward to continued work by the Methods Working Group in the future.

Oscar A. Garcia

UNEG Chair

New York, December 2022

Introduction

The period 2021-2022 was characterized by the deepening of the COVID-19 pandemic along with the advent of an era of ‘polycrisis’: The climate emergency is hitting harder and faster than projected; global inequalities have risen to early 20th century levels; conflicts are at their highest levels since the Second World War, exacerbating a cost-of-living crisis and rising food and energy insecurity. As the world’s population reached 8 billion in November 2022, the number of forcibly displaced people passed the 100 million mark for the first time ever.

The evaluation community has a significant role to play in shaping the choices and pathways that could be taken to pivot into the future we want. In particular, the exponential rate of innovation and technological advances in all parts of the world are bearing great promise in our ability to harness humanity’s immense collective power to tackle the defining challenges of our times.

As evaluators, we have stayed and delivered throughout the worst of COVID-19. Evaluations were conducted remotely, drawing more heavily on secondary sources and online interviews, and increasingly relying on the talent and contextual knowhow of national and local evaluators. At various points, the evaluation community reflected on ways to capture and amplify the voices of the most left behind populations – those not easily reached by remote interviews. It learned lessons from past crises and from those regularly evaluating under crisis conditions, adjusting evaluation criteria, questions and methods. And it questioned the need to fly international evaluators around the world, adding to the collective carbon footprint of our profession.

In this context, the work of the UNEG working group on methods was more pertinent than ever. Through its workstream on ‘use and appropriateness’ of methods, UN evaluation practitioners continued to reflect on appropriate evaluation designs and discuss what adjustments worked and didn’t work as the crises evolved. Early in 2021, the group held two webinars, on “*Version Two: [Choosing Appropriate Evaluation Methods – A Tool for Assessment and Selection](#)” and “*Adapting Evaluation during the COVID-19 Pandemic: the [Rainbow Framework](#)”* to set the scene. Both approaches emphasize the need to select evaluation methods suitable to the questions the evaluator wishes to answer, and feasible given evaluability considerations and the monitoring data at hand. The Rainbow Framework also assists with systematically planning and implementing evaluation activities from start to finish. Working group members discussed how they used the two frameworks for revising their workplans, re-designing their on-going evaluations and designing new evaluations in the context of COVID-19 restrictions.*

Through short and concise summaries of evaluation methods, it is our hope that this compendium will encourage evaluation practitioners to further broaden and sharpen their methodological toolbox in order to produce evaluations that are useful in changing times, to support the UN family in maintaining international peace and security and promote social progress, better living standards and human rights.

The UNEG Methods Group coordinators

The Methods Group Members

#	Agency	Name
1	UNHCR	Henri van den Idsert
2	UNHCR	Francesca Carini
3	UNAIDS	Elisabetta Pegurri
4	UNFPA	Alexandra Chambel
5	UNFPA	Valeria Carou Jones
6	GEF	Anupam Anand
7	UNESCO	Ahmedou El Bah
8	UNESCO	Taipei Dlamini
9	IOM	Jhonn Rey
10	IOM	Angeline Wambanda
11	IOM	Zahoor Ahmad
12	IOM	Abderrahim El Moulat
13	IOM	Martin Schmitt
14	IOM	Sarah Lynn Harris
15	ILO	Peter E. Wichmand
16	OIOS	Pankaj Verma
17	OIOS	Hanife Cakici
18	FAO	Anshuman Bhargava
19	FAO	Carlos Tarazona
20	UNRWA	Leslie Thomas
21	UNDP	Tina Tordjman-Nebe
22	UNDP	Harvey Garcia
23	WFP	Felipe Dunsch
24	WFP	Simone Lombardini
25	WFP	Dawit Habtemariam
26	UNEP	Janet Wildish
27	UNICEF	Kathleen Letshabo
28	UNICEF	Simon Bettighofer
29	UNICEF	Eduard Bonet
30	UNICEF	Kamilla Nabiyeva
31	UNICEF	Zlata Bruckauf
32	UNICEF	Adrian Shikwe
33	UNICEF	Xin Xin Yang
34	UNICEF	Carlos Rodriguez-Ariza
35	UNICEF	Sultanat Rasulova

Setting Our Priorities

In defining our focus of work for the period 2021-22, and remembering our primary objective for this work to be useful to evaluation practitioners working across the full range of UN system work, the group came up with the following priority areas:

1. Methods and approaches that facilitate real time learning.
2. The use of technology and social media for [remote] data collection.
3. How best to support full participation of partners and beneficiaries, bearing in mind the centrality of the *Leaving No One Behind* objectives to the SDGs and the challenges posed by the pandemic to reach out the most vulnerable groups.
4. Methods and approaches to assess the causal relationship between what UN agencies deliver and the results that they claim at different levels of their impact pathways.

Methods Discussed

The table below groups the sessions organised under the above four overarching areas; each of them is discussed in a separate brief which presents the main features of the methods and approaches reviewed as well as any challenges and lessons learned that emerged from their application.

Table 1: Themes Reviewed by the Working Group (2021-2022)

#	Method	UN agency and guest speakers
1.	Real-time learning approaches (Developmental Evaluation)	UNFPA, WFP, Inspire to Change (Dr Nora Murphy Johnson)
2.	Use of technology and social media for evaluation data collection (Geospatial Methods and Social Listening)	GEF, MINUSMA, UNICEF, London School of Hygiene and Tropical Medicine (Dr Heidi J. Larson)
3.	Stakeholder engagement in evaluation	FAO, UCL Institute of Education (Dr Sandy Oliver)
4.a	Advanced analytical techniques that can be applied remotely (Process Tracing)	UNDP, University of Aarhus (Dr Derek Beach), ITAD (Dr Zoe Sutherland)
4.b	Case study and case-based approaches in evaluation	Dr Linda Morra Imas, independent consultant

1. Real-time learning approaches

1.1. Developmental Evaluation

This brief describes the use of **Developmental Evaluation**, its key features as well as lessons learnt from applying this evaluation approach within the UN context. Experiences discussed during the two webinars on Developmental Evaluation organised by the UNEG Evaluation Methods Working Group in 2021 and 2022 are also presented.¹

Why conduct a developmental evaluation?

Developmental Evaluation (DE) is a term coined by Michael Quinn Patton in the mid-1990s. This approach has gained more and more traction over the past years as being particularly suited for assessing, informing and supporting innovative and adaptive development, especially in complex and dynamic environments where a knowledge base is not yet established. DE can be used for a range of purposes, including ongoing programme development, generating innovations and taking them to scale, facilitating rapid response in crisis situations and accompanying organizational change.²

Key characteristics of DE?

- Focus on development and adaptation (versus improvement, accountability or summative judgment)
- Takes place in unpredictable complex dynamic environments
- Feedback is rapid, or as real-time as possible
- The evaluator works co-creatively with social innovators to conceptualize, design and test new approaches in a long-term, on-going process of adaptation

What are the advantages of this method?

DE is **highly adaptive and versatile**. It helps identify options and solutions while navigating through non-linear and complex environments, for which more traditional evaluation approaches are not always well suited.

It is based on a **forward-looking and utilization-oriented approach**. DE involves multiple co-creation processes with evaluation users, thus promoting a high level of participation and ownership of results.

It entails a continuous process of **data collection and feedback loops**, which support timely decision-making and adaptation. This real-time feature of DE helps maximise utility and learning while also enabling quick course correction.

What are the disadvantages of this method?

DE implies a high-level of flexibility, as it is **more time consuming and resource-intensive** than traditional evaluation approaches. Budgeting for a DE can be difficult since this does not entail a predictable process from the onset.

¹ The first webinar took place on November 10th 2021, and featured presentations from the World Food Programme (WFP) and the United Nations Population Fund (UNFPA). The second webinar took place on June 23rd 2022 and featured an expert lecturer (Dr Nora Murphy).

² Michael Quinn Patton, *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use* (Guilford Press, 2011)

Compared to a summative evaluation approach (and most formative approaches), DE entails a **higher level of stakeholder engagement** through a regular sequence of feedback loops and co-creation activities as well as an iterative process of data collection. If not carefully planned, this high level of participation could cause fatigue while the large amount of data generated may be difficult to absorb.

When should you use/not use this method for evaluation?

DE is not necessarily appropriate to all contexts and evaluation purposes. DE is particularly useful when there is willingness to experiment, there are no predefined theories of change or known solutions to a given problem, and multiple pathways are possible.³

Different from formative evaluations, whose primary purpose is to gather information that can be used to improve or strengthen the implementation of already established models or programs, DE is more suited for developing new initiatives or also testing new models.⁴

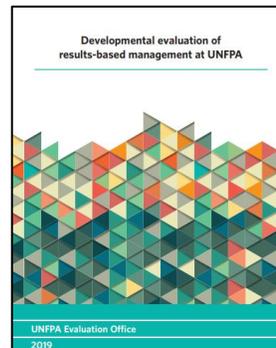
In case of doubt, a fidelity check can be run to understand whether DE is the most appropriate method to use.⁵ Hybrid models, combining DE principles with elements from other evaluation approaches, can also be explored (e.g., see WFP’s experience below).

How has this method been used in the UN?

UNFPA: In 2019, the UNFPA Evaluation Office conducted a centralised DE, to develop the next stage of its results-based management (RBM) system. DE was used to help identify the root

causes of persistent bottlenecks that hindered the effectiveness of the old RBM, while also bringing in new evidence and solutions for the way forward.

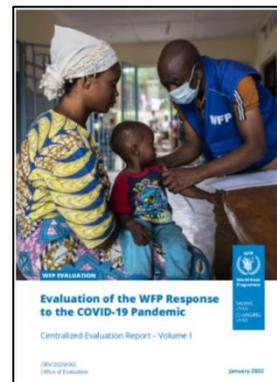
Starting from the identification of five creative tensions (gaps and challenges), which prevented the old RBM system from optimizing its performance, an evaluative inquiry framework was outlined. Through



different data collection and feedback loops, six leverage points (areas where small changes can produce large improvements) were identified as possible solutions. These included the development of a shared conceptual framework on RBM; revising the RBM system requirements, procedures and tools; increasing the use of evaluations; revamping HR competency and recruitment frameworks; looking into behavioural transformation; and increasing dialogue with the Executive Board on accountability and reporting.

See [here](#) for additional information.

WFP: Between February 2020 and June 2021, WFP conducted an independent evaluation of its response to COVID-19 to meet both the organisation’s learning and accountability needs. The evaluation adopted a “retrospective



³ Ibid.

⁴ Patton, M. Q. (2009). Developmental evaluation as alternative to formative assessment. [Web Video available [here](#)].

⁵ See an example of a diagnostic checklist for DE [here](#).

developmental” design, which focused on examining the adaptive capacity of WFP under pandemic conditions. This was a hybrid approach, which combined DE elements with a summative component on assessing results. Specific emphasis was placed on how the response developed, rather than a traditional theory-based evaluation. Cross-fertilization, stakeholder engagement, and regular feedback loops were central to this approach.

See [here](#) for additional information.

What challenges were encountered?

- For many, DE is still a new concept that requires a shift in mindset, especially for those used to working only with summative evaluation approaches. The flexible nature of DE as well as the novelty of some of its principles can lead to confusion, for both evaluators and evaluands.
- Reluctance to engaging in co-creative and highly participatory processes remains a common challenge in the practice of this evaluation approach.
- Evaluators’ independence may be questioned, as these are typically embedded within the same team that is subject to the DE. This means that a new type of relationship between evaluators and evaluands has to be established.
- Conducting a DE requires a different set of skills and competencies compared to summative evaluations. Finding experienced consultants who can undertake this type of exercise in the UN system can often be time consuming.
- The UN is still very much accountability-focused when it comes to evaluation, while DE is more learning- and utilization-oriented. Finding a balance between both learning and accountability needs within the UN can be challenging.

- In entities where programmes are fully established with a standard and rigid theory of change or/and approach for implementation, incorporation of or adherence to DE principles may be difficult at this stage of implementation.

Key lessons learnt

- Holding a workshop with stakeholders early on in the DE process to explain what this approach entails and what its benefits are, can be highly beneficial in mitigating any negative perceptions. This ensures that those who are new to DE are aware of its key features and added value. Such workshops can also help foster engagement, build trust and lay down a solid ground for collaboration.
- Managing expectations as early as possible is essential to avoid disappointment. For example, conducting a DE in a large-scale emergency response is not always feasible, so risks and limitations should be clear from the onset. Scoping exercises are recommended to better understand whether a DE approach is appropriate for a given context or if other options (including hybrid ones) should be pursued instead.
- DE requires a substantial investment in both human and financial resources. It also means devoting time to process management, which can be time consuming and place a big burden on the commissioning entity. A mitigation measure would be to outline a detailed DE process well in advance.
- Conducting DE quality assurance can be challenging. In absence of tools available for this purpose, the UNFPA Evaluation Office decided to develop its own developmental evaluation-specific QA grid and guidance.⁶ This turned out to be very useful in assessing the quality of DE.

⁶ See UNFPA QA grid for DE [here](#)

Resources

Books

Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: Guilford Press.

Gable J. et al. (2021). *A Development Evaluation Companion*. The McConnell Foundation.

Links

Assessing the Quality of Developmental Evaluations in UNFPA. Link [here](#).

Better Evaluation Website. *Developmental Evaluations*. Link [here](#).

Blue Marble Evaluation Initiative. Link [here](#).

Developmental Evaluation of Results-Based Management at UNFPA. UNFPA Evaluation Office (2019). Link [here](#).

Diagnostic Checklist for Developmental Evaluation by Cabaj M. Link [here](#).

Evaluation of WFP's Response to the COVID-19 Pandemic. WFP Office of Evaluation (2022). Link [here](#).

2. Use of technology and social media for evaluation data collection

2.1. Geospatial methods

This brief describes the use of geospatial methods in evaluations and social listening as a method applied to inform programme design. Experiences, challenges and lessons learned from applying these methods within the UN context also shared in this section. Insights and reflections emerged during the webinars on remote sensing and social listening organised by the Group in 2021 are also included.⁷

What is a geospatial method?

Geospatial data is unique as it has spatially explicit information. It can be gathered from remote sensing platforms, geotagged photographs, and ground sensors. Survey datasets also have geolocation information. The geospatial methods include the creation, collection, analysis, visualization, and interpretation of geospatial data.

Geospatial data and methods, therefore, can provide spatially explicit, synoptic, time-series data for various earth system processes, and therefore have been used in the monitoring and assessment of environmental processes for the past four decades.

What are the advantages of using geospatial approaches in evaluation?

- Geospatial methods are efficient. We don't need to visit all the intervention sites. It can save us financial and human resources and can be very useful when working in hard-to-reach areas, especially in fragile and conflict-affected situations (FCS).

- Geospatial analyses are scalable, and one can perform analyses at the project site, portfolio, or global level.
- These methods provide objective evidence and thereby aid transparency.
- And these approaches work well in a mixed-methods framework and help deal with common evaluation challenges such as lack of baseline, finding the right counterfactuals, and addressing accessibility issues.

What are the disadvantages of this method?

- Depending on the scale and purpose of the evaluation, geospatial data and associated analyses may require high-performance computing generally not available in evaluation offices, and could be expensive.
- Availability and accuracy of contextual variables are often an issue and vary widely across countries and sites.
- These approaches also require field verification.

When should you not use this method for evaluation?

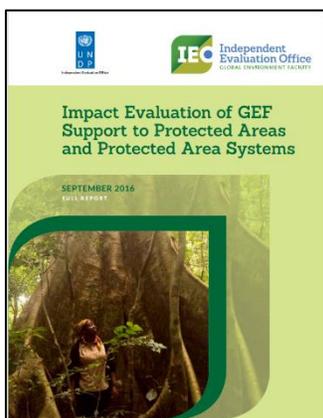
⁷ The two webinars were organised in 2021 and featured speakers from GEF, MINUSM, UNICEF, and the London School of Hygiene and Tropical Medicine.

Despite its usefulness in answering a variety of evaluation questions, this method may not be appropriate for answering “how” and “why” questions.

How has this method been used in the UN?

“Evaluation on Impact of GEF Support to Protected Areas” jointly conducted by the Independent Evaluation Offices (IEOs) of the GEF and UNDP. The evaluation assessed the impact of GEF investments in non-marine protected areas (PAs) and PA systems.

As the financial mechanism for the U.N. Convention on Biological Diversity (CBD), the GEF's Strategy is consistent with the CBD's Strategic Plan and is reflected by its support to protected areas



for the past three decades. Between 1991 and 2015, the GEF provided \$3.4 billion in grants to 618 projects, matched by \$12.0 billion in co-financing, to help protect almost 2.8 million km² of the world's non-marine ecosystems (GEFIEO 2016).

Assessing the effectiveness and impact of GEF supported protected areas was challenging mainly due to the scale, different timelines, and difficulty in collecting primary data due to the remoteness of protected areas. These challenges were addressed by using geospatial approaches.

At the global level, the evaluation used many observations from satellite data, and geospatial analysis covering 2001-2012 in GEF supported protected areas and their buffers at 10 and 25 km to compare the extent of forest loss between

these areas. The global analysis component measured outcomes using forest cover (geospatial analysis of 580 PAs in 73 countries), wildlife populations (88 species in 39 PAs), and Management Effectiveness Tracking Tool (METT) scores (2440 METTs from 1924 PAs in 104 countries) as indicators. The portfolio analysis component included a total of 618 projects in 137 countries.

The evaluation concluded that GEF support contributed to biodiversity conservation by helping to lower habitat loss in PAs and to build capacities that address key factors affecting biodiversity conservation in PAs. Sustainable financing of PAs was identified as a concern.

What challenges were encountered?

The key challenges encountered during the evaluation were: *a. Substantial information gaps on GEF support to PAs b. limited global time-series data available for GEF supported PAs, and c. difficulties in estimating counterfactuals, or what would have happened without GEF support.* The evaluation dealt with information gaps by pooling resources and sharing management of the evaluation, collecting data from multiple sources including field interviews. The limitations posed by global time-series data was mitigated by performing several types of data analyses. For counterfactual analysis, and to increase comparability and minimize the overestimation of GEF's impact, GEF PAs were compared only with non-GEF PAs within the same biomes located in the same countries. The evaluation used a mix of quantitative, qualitative and spatial methods in data collection and analysis to address the gaps and systematic biases in the datasets. Evaluative evidence was also collected from a mix of sources, combining global datasets, field data, literature reviews, and statistical models.

Key lessons learnt

Geospatial approaches help generate evaluative evidence that is not possible through traditional evaluation methods. The satellite driven spatial analysis enabled assessment of the effectiveness of more than 500 GEF-supported protected areas in a timely and cost-efficient manner.

Geospatial approaches require specialized technical skills and multidisciplinary teams to conduct complex evaluations. Successful implementation needs a suite of tools and multidisciplinary teams. It is, therefore, essential to work with a multidisciplinary team right from the planning stage for the assessment through to its execution, analysis, and interpretation of evaluative evidence.

2.2. Social Listening

What is social listening and what are the advantages and disadvantages of this method?

Social listening is a process aimed at analysing different online sources including blogs, social media platforms (e.g., Facebook, Twitter) and forums, to gain insights on how a given population of interest thinks and behaves about a certain topic. For example, UNICEF has used social listening to measure the influence of social media on the decreasing immunisation rates in selected countries, as well as to assess their impact on attitudes, trust and immunisation intention (e.g., for Covid-19).⁸ The evidence collected was then used by the organisation to support decision-making and programming in the area of demand creation and vaccine acceptance. This example shows how social

listening can help producing a content that understands and resonates well with our target audiences. It also provides rapid and real time insights on sentiments and narratives among a large set of audiences, while highlighting topics of interest, misconceptions and rumours. This method, however, also comes with some challenges, such as finding and filtering information in crowded and ever-changing online spaces, performing sentiment analysis or assessing the quality and reliability of the data collected for our research purposes.

Application in the UN

UNICEF with the help of the London School of Hygiene and Tropical Medicine sought to measure the influence of social media on the decreasing rates of immunization in selected countries. In addition, the Regional Initiative on Social Listening in Europe and Central Asia sought to assess the impact of social media and anti-immunisation movements (including Covid-19) on attitude, trust and immunisation intention. This evidence was used to support decision-making in programming in the area of demand creation and strengthening vaccine acceptance. The ultimate aim was to create a toolkit to drive smart decisions around content, targeting and audience prioritization.

Core pillars of social media listening



⁸ Presentation given by UNICEF in February 2021 led by Professor Heidi Larson, London School of Hygiene and Tropical Medicine.

The main social listening platforms analysed were Facebook, Twitter, Instagram, Youtube, Ok.ru, and Vk.ru. The main social listening tools were Meltwater (listening to global conversations), Facebook analytics, Talkwater (automated social listening reports/dashboards), Youscan, and SoTrender. The vast amount of data used to analyse national and regional sentiments disaggregated for certain population groups and demographics, allowed UNICEF to develop targeted campaigns on the same social media platforms. The tools at their disposal also allowed UNICEF to monitor and assess the success rates of their campaigns through key metrics (sentiment changes, views, clicks, sign-ups), as well as monitor changes in online conversations and topics.

[Tracking anti-vaccination sentiment in Eastern European social media networks \(2013\) UNICEF Europe and Central Asia](#)

Resources

Books

Anand and Batra (2021). Using big data and geospatial approaches in evaluating environmental interventions. In Uitto, Evaluating Environment in International Development

Papers

M Lech, JI Uitto, S Harten, G Batra, A Anand (2018). Improving international development evaluation through geospatial data and analysis. International Journal of Geospatial and Environmental Research 5 (2), 3

Reports

GEF IEO (2015). [Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems](#)

[UNICEF ESARO: COVID-19 and its impacts \(2021\)](#). Social listening report on online conversations in Eastern and Southern Africa

3. Stakeholder engagement in evaluation

[The brief is currently being finalized, and will be integrated in full in February 2023]

4a. Advanced analytical techniques that can be applied remotely (Process Tracing)

This brief describes the use and key features of Process Tracing. This method can enrich the evaluation toolbox of UN entities seeking to trace how and why an outcome came about by unpacking the causal mechanism at play. Its potential unfolds if applied to complex impact patterns with well definable theories of change, such as policy influence. Insights and reflections from three different webinars on process tracing that were organized by the Group in 2021 are included.⁹

What is Process Tracing?

Process Tracing (PT) is a theory-based evaluation method to assess the outcome of an intervention through unpacking the underlining causal mechanisms. Observed evidence is meticulously linked to hypotheses in order to reduce biases and shed light on the contribution of various factors to the outcome. Four formal tests assess the type and strength of the evidence, which can be supported by applying Bayesian principles¹⁰. Those tests are guided by the questions “how probable is the observation of the piece of evidence if the hypothesis is true” (sensitivity) and “how probable is the observation of the hypothesis is false” (type 1 error).

What are the advantages of this method?

- Rigorous within-case methodology allowing to quantify qualitative data.
- In-depth understanding of mechanisms that caused change, providing rich information on how an intervention worked.
- Offers a tool to challenge collective and evaluative biases for its “forensic” character, sometimes linked with detective metaphors.

- Backwards from observed outcomes to potential causes to infer causality, which allows to evaluate past non-linear interventions implemented overtime.
- Tests multiple theories of causality-in-action, which helps finding counterfactual explanations
- Provides evidence with different kinds of probative value.

What are the disadvantages of this method?

- Granular data is required to assess each contribution claim, akin to mini evaluations, in-depth case studies or interventions with small sample sizes.
- Within-case methodology doesn’t lend itself to multi-level or multi country evaluation questions.
- Application of the Bayesian updating formula requires a solid methodological understanding.
- Requires detailed, explicit and empirical theorization of expected causal relations and their inferential underpinnings, which goes far beyond a regular theory of change.

When should you use/not use this method for evaluation?

⁹ The three webinars on Process Tracing featured presentations from UNDP, Prof. Derek Beach (University of Aarhus), and Ms Zoe Sutherland (Senior Consultant at ITAD).

PT is a powerful tool if the evaluation team has a neat contribution hypothesis, a detailed and empirical theory of change, and is confident that the evaluand can and will provide highly granular data (e.g., emails, report versions with track changes, minutes from meetings etc.).

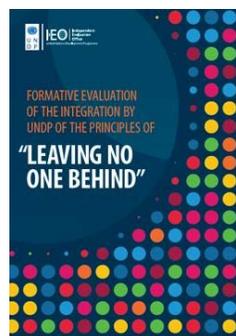
The application is particularly valuable if the influence of a certain intervention or flagship product is frequently claimed but in-depth evidence is missing, impacts are intangible and/or counterfactual explanations cannot be found. In this sense, PT is particularly recommendable to evaluate interventions aimed at influencing policy processes.

PT should **not be used** if the evaluation team:

- Doesn't have a precise hypothesis, and detailed and empirical theory of change.
- Doesn't assess a high-profile claim, intangible impacts or flagship intervention. Otherwise, the workload may not be justified.
- Has a mainly comparative interest or would like to trace impact pathways across levels (e.g., country and global).
- Has doubts about the evaluands' ability and willingness to provide (internal) data.

How has this method been used in the UN?

UNDP applied process tracing in its formative evaluation on the integration of 'leave no one behind' principle to substantiate the claim "UNDP helps improve livelihoods for the most deprived by supporting countries in determining who is 'left behind'".



The analysis fed into chapter 5, using UNDP's work on data ecosystems in five country contexts as an entry point.

What challenges were encountered?

PT was integrated into a corporate evaluation conducted at a central evaluation unit as part of its executive board requirements, which requires a degree of aggregation to inform strategic decision making. However, the strength of PT lies in with-case analysis and subsequently possibilities for horizontal (e.g., across countries) and vertical aggregation was limited, or would have appeared forced. Defining one main outcome hypothesis was a challenge, especially as UNDP aimed at tracing the same overarching claim in several countries.

The formal assessment of each mechanism strengthens rigour yet requires extensive data and almost forensic diligence. To unpack the 'how' the evaluand has to provide copious amounts of (internal) information. Even committed teams struggled to provide the required level of support to the evaluation. Obtaining internal documents (e.g., email chains, observations) was even more challenging if the intervention took place some time ago due to staff turnover, etc.

In some instances, "proving" the contribution of a UN entity also ran counter to fostering national ownership among stakeholders, e.g. in this evaluation in the case of technical assistance provided to LGBTIQ+ sensitive legislation.

Resources

Andrew Anguko (2019): [‘Process Tracing as a Methodology for Evaluating Small Sample Size Interventions’](#). African Development Bank, Evaluation Matters Second Quarter.

CDI (2015): [‘Applying Process Tracing in Five Steps’](#). Centre for Development Impact. Practice Paper Annex 10.

Barbara Befani (2021): [‘Credible Explanations of Development Outcomes: Improving Quality and Rigour with Bayesian Theory-Based Evaluation’](#). Report 2021:03 to The Expert Group for Aid Studies (EBA).

Benedict Wauters and Derek Beach (2018): [‘Process tracing and congruence analysis to support theory-based impact evaluation’](#). SAGE, Vol. 24(3) 284-305=

Jacob Ricks and Amy Liu (2018): [‘Process Tracing Research Designs a Practical Guideline’](#). Political Sciences and Politics, Vol 51:4, p.842- 846.

Johannes Schmitt and Derek Beach (2015): [‘The contribution of process tracing to theory-based evaluations of complex aid instruments’](#). SAGE, Vol. 21(4) 429-447

UNDP IEO (2022): [Formative Evaluation of the Integration by UNDP of the Principles of ‘Leaving no one Behind’](#). United Nations Development Programme, New York.

4b. Case study and case-based evaluations

This brief is divided in two parts. The first part provides an overview on case studies and case-based approaches in evaluation. The second part (*to be completed in February 2023*) will include lessons learnt from applying this evaluation approach within the UN context.

Why conduct a case study?

“A case study is an empirical inquiry that investigates a phenomenon in depth and within its real-world context, especially when the phenomenon may be context-dependent”¹¹. UN agencies resort to this technique to analyse complex interventions and draw lessons from these. Depending on the purpose of each evaluation, the case study may delve into one or several specific aspects of an initiative in order to identify contributing factors for success and failure or to compare different implementation contexts. It is a particularly effective method to achieve conceptual validity, to examine the assumptions around causal mechanisms and to assess causal complexity¹².

Key characteristics of a case study

- A case study provides a holistic view of a complex intervention, often detailing the context and the intervention itself.
- It focuses on a specific intervention and may cover one or several implementation settings for this intervention (single case or multiple cases).
- It is methods neutral and often uses both qualitative and quantitative methods. Evaluators must triangulate different data sources to analyse the data drawn from case studies.

- It can be used to answer descriptive, causal and evaluative questions within a broader evaluation framework.
- It allows to test theories of change and identify patterns.

What are the types of case studies?

- **Explanatory case study:** It seeks to test a theory of change and the causal links between the project and the supposed results achieved. As it seeks to answer the questions ‘why’ and ‘how’, it is particularly fitting to explain the factors that have contributed to an intervention’s success or failure.
- **Descriptive/Illustrative case study:** It describes the context in which an intervention is embedded and the results it has generated. This is often a factual and objective description of the implementation of a project, with little additional analysis. It is useful to understand the achievements of a project in a given context and to provide concrete examples of programme implementation.
- **Exploratory case study:** its purpose is to ascertain the potential outcomes of an intervention by looking into specific examples of implementation. It is particularly used when there is limited available data on an intervention’s outcome and the evaluator seeks to establish a pattern before fully designing his evaluation framework and questions.

¹¹ Yin, R.K. (1989) Case Study Research: Design and Methods. Sage Publications, Newbury Park.

¹² George, A., & Bennett, A. (2005). Case studies and theory development in the social sciences. Cambridge: MIT Press.

What are the advantages of this method?

Case studies provide an opportunity to assess and at times compare an intervention undertaken in different contexts and more clearly **identify and outline results**. Case studies are thus particularly effective to **identify good practices and lessons learned**. They also help identify the **unintended consequences** of an intervention and examine causal mechanisms and/or identify what contextual conditions activate causal mechanisms.

Case studies enable the evaluator and the audience to understand key underlying mechanisms in programme implementation and programme success through concrete **in-depth examples, that are easy to communicate and grasp**. Case studies are stronger at examining if and how a variable affected the outcome rather than assessing how much it affected it (see limitations), and at testing assumptions on whether a variable is a necessary or sufficient condition for a certain outcome.

They use a **variety of data sources** (key informant interviews, focus group discussions, desk review, observation, etc.). It is thus a complex evaluative piece in and of itself.

As for other approaches, the case study methodology **engages a variety of stakeholders**, contributing to greater ownership and understanding of the intervention, namely by involving other stakeholders besides programme managers in the evaluation process. Including them also in feedback loops further enhances this.

What are the disadvantages of this method?

Without a clearly defined evaluation framework and objectives, case studies may turn into a weak evaluation tool. One of the common limitations of case studies is case selection bias. It requires rigour and **prior scoping** for it to yield results. It is essential to identify specific questions in order for the case study to feed into the overall evaluation framework.

In some evaluations, case studies are rather case examples than a research method when a certain country or intervention is described in detail.

Oftentimes, in the UN context, a case study is **heavily interview-based** as the evaluator seeks to gather the reflections of all stakeholders involved in an intervention ranging from programme implementers to beneficiaries and partners. In these situations, case studies become too heavily reliant on individual opinions. This entails a **higher risk of bias** and subjectivity in the analysis, as interviewees' interpretations and recollection of events may vary, and they may choose to withhold more critical information if this is perceived as potentially jeopardizing the continuation of an intervention. To triangulate and verify the data thus requires not only sampling a larger number and a variety of stakeholders – which may prove **more time consuming** for the evaluation team – but also weighting the views of some stakeholders compared to others, which may introduce **some other level of bias** in the data analysis. As in any evaluation, it is also essential to complement the analysis with other sources (e.g. desk review, surveys). Whilst the case study allows evaluators **to establish some causal links between the intervention and the results observed it remains challenging to attribute impact to a sole intervention**. Furthermore, the strength of the inference will vary significantly depending on the solidity of the existing theory of change, the quality and scale of evidence collected and the evaluators' capacity to test each assumption and eliminate other contributing factors. It is crucial to also consider and acknowledge the role of other players and

projects in complementing or hindering the achievement of one own objectives.

It is key to also underline the **limited generalizability of a case study**. Case studies cannot extrapolate the finding and guarantee that the application of the same methods that proved successful in one case study would yield the same results in another setting; unless the evaluators identify similar patterns across different contexts, identified both through extensive literature review and several heterogenous cases for similar interventions. This is oftentimes difficult to achieve within resource and time constraints for evaluations. Hence, caution must be applied when generalizing the findings.

When should you use this method for evaluation?

Case studies are versatile and can be used to fulfil several evaluation needs.

This method is particularly useful when seeking to identify good practices to emulate and poorer mechanisms to avoid; in order to learn from the practical implementation of a project in a specific context. They are ideal to test assumptions and understand facts. For example, they are interesting when looking at cases where good results were yielded while in others no results were achieved (good and bad examples).

They are also ideal when seeking to collect qualitative data to explore the underpinnings of an intervention and provide more context to justify the observations made in an evaluation.

How has this method been used in the UN?

[Section to be completed in February, after the second part of the webinar on case studies takes place]

What challenges were encountered?

[Section to be completed in February, after the second part of the webinar on case studies takes place]

Key lessons learnt

[Section to be completed in February, after the second part of the webinar on case studies takes place]

Resources

Books

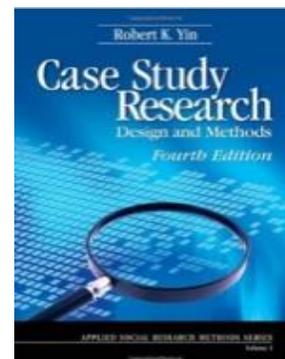
Morra Imas, Linda and Rist, Ray (2009). The Road to Results. Designing and Conducting Effective Development Evaluations.

Stake, Robert E. (1995). The Art of Case Study Research. California, Sage Publication INC;

Yin, Robert K. (2017). Case Study Research and Applications: Design and Method. California, Sage Publications Inc.

Scholz, Roland and Tietje, Olaf (2002)

Embedded case study methods (Integrating quantitative and qualitative knowledge), California, Sage Publications Inc.



Links

Morra, Linda and Friedlander, Amy (1995). World Bank Operations Evaluation Department. Link [here](#).

Better Evaluation Website. *Case study*. Link [here](#).

Better Evaluation Website. *Comparative case studies*. Link [here](#).

Concluding Remarks

The value added of UNEG as a learning community has been demonstrated in the work of the methods Working Group – first constituted in 2019 – via a series of useful discussions on approaches, techniques and methods that are relevant to evaluate the work of the UN system at different levels (community, national or global).

We can think of the work of this group as a contribution to UNEG professionalization in two ways: first, in its attempt to formulate and begin to populate a universe of different evaluation designs that is of relevance to an audience of UN evaluators, drawing on good practice in evaluation and social science research more generally; second, in its attempt to begin to document evidence of the application of these designs across a number of areas of the work of the United Nations.

Workplan objectives 2023

The Working Group aims to continue with the peer-to-peer learning carried out to date. Given that there are many methods of potential utility that have not been discussed yet, there is a need to establish some criteria for determining future priorities and/or expanding into new areas of work.

For instance, the group could contribute to broader discussions regarding the development or updating of methodological guidance for UN evaluations at the thematic, policy, strategy, and programme/project level.

By sharing experiences from UN agencies, external partners and evaluation's communities of practice, the group at its core will continue to help evaluation practitioners navigate the universe of methods in a useful and structured manner – in order to tailor them to their needs.

The course and focus of the Working Group will be set following discussions at 2023 UNEG AGM.

Speakers' short bios

Anupam Anand: Anupam Anand is a Senior Evaluation officer at the GEF IEO. He has more than 15 years of combined experience in evaluation, international development, academic research, and teaching. Anupam has led evaluations on biodiversity, SFM and REDD+, land degradation, fragility and conflict, and illegal wildlife trade. He uses a blend of innovative approaches such as satellite data, GIS, machine learning, drones, and field-based methods to enhance evaluative evidence. Before GEF, he was a Remote Sensing Scientist at the Global Land Cover Facility (GLCF), working on multiple NASA funded projects, including field campaigns for future satellite missions, and also worked for the Climate Investment Funds. He has published several scientific articles and book chapters. Anupam holds a Ph.D. on applications of lidar remote sensing from the University of Maryland.

Barbara Befani: Dr. Barbara Befani is affiliated with the University of Surrey and the University of East Anglia but working mainly as an independent researcher / consultant. Her interests include 1) evaluation quality; 2) methodological appropriateness and comparative advantages and weaknesses of different evaluation methods; 3) causal inference frameworks for impact evaluation; and 4) specific hybrid, qualitative methodologies, like Qualitative Comparative Analysis (QCA) and Process Tracing (in particular its Bayesian formalisation, which she is extending to all forms of Theory-Based Evaluation). She's a former Secretary General of the European Evaluation Society.

Derek Beach: Derek Beach is a professor of Political Science at Aarhus University, Denmark, where he teaches European integration and research methodology. He has authored articles, chapters, and books on research methodology, policy evaluation, international negotiations, referendums, and European integration, and co-authored the book *Process-tracing Methods: Foundations and Guidelines* (2019, 2nd edition, University of Michigan Press). He has taught case study methods at numerous workshops and ph.d. level courses throughout the world, and conducted evaluations at the national and international level. He was an academic fellow at the World Bank's Independent Evaluation Group in spring 2022. He is an academic coordinator of the Methods Excellence Network (MethodsNet).

Julia Betts: Dr Julia Betts specialises in the evaluation of international assistance in conflict-affected and other operationally-challenging settings. She has designed and led many complex transnational evaluations for bilateral and multilateral agencies, with a focus on systematic methodologies in data-poor environments. Julia supported the OECD DAC in the adaptation of the international evaluation criteria 2018-2020; and led WFP's 2020-2021 large-scale developmental evaluation of its COVID-19 humanitarian response.

Jordi del Bas Avellaneda: Jordi is a professor of Global Corporate Strategy at the Maastricht School of Management (MSM), which is part of the School of Business and Economics (SBE) of Maastricht University. He is a senior accredited trainer on evaluation by the Secretariat General of the European Commission and at the European Institute of Public Administration (EIPA). Besides teaching and conducting academic research, Jordi works as an independent evaluator and organizational development specialist.

Hanife Cakici: Hanife, PhD in Evaluation Studies, is an Evaluation Officer at the Peacekeeping Evaluation Section of the Office of Internal Oversight Services (OIOS). Prior to joining the United Nations, Hanife has evaluated several public policy efforts in the U.S., Europe, Middle East, and Sub-Saharan Africa. Her most recent evaluations included transition of the United Nations from peacekeeping to peacebuilding in Sudan, MINUSMA's contribution to rule of law in Northern and Central Mali, and Women, Peace and Security in elections and political transitions, and UNMIK contribution to rule of law. In addition to OIOS,

she worked at the Regional Economic Commission for Africa and the Department of Economic and Social Affairs. Hanife served as a gender focal point for several years, working closely with UN Women on UN-SWAP framework and methodology.

Valeria Carou-Jones: Valeria Carou Jones is an Evaluation Adviser at UNFPA Evaluation Office since 2011 and conducts corporate evaluations in the areas of adolescents and youth, gender, maternal health and institutional level exercises on innovation, south-south cooperation, and results-based management. Her past experiences include working as a consultant for 9 years on evaluations for UNDP, the European Commission and USAID. She has also worked at the World Bank for 4 years in the Infrastructure and Private Sector Development Unit. She has an MA in international economics from Johns Hopkins.

Josep Maria Coll Morell: Josep M. Coll, PhD is an economist specialized in international and sustainable development. He works as an independent evaluation and organizational development consultant for a wide range of public and private organisations, exploring the frontiers of new utilization-focused evaluation approaches that fit complex and turbulent contexts. He was a member of the team that conducted the first corporate developmental evaluation (in UNFPA) ever conducted in the United Nations system.

Taipei Dlamini: Taipei Dlamini is an Evaluation Specialist at the UNESCO Division for Internal Oversight Services. Over the past five years, she has evaluated a variety of UNESCO programmes including the evaluation of UNESCO's work on Media and Information Literacy, its action to protect culture in emergencies or that of UNESCO's Strategy for action on Climate Change. She holds a Master's degree in International Law and Administration from the Université Paris I Panthéon Sorbonne. Prior to joining UNESCO, she worked briefly with the Southern African Development Community (SADC) Lawyer's Association, a South African-based NGO aimed at promoting the rule of law and the protection of human rights in the SADC region.

Solveig Gleser: Solveig Gleser is an evaluation specialist with a focus on evaluations in contexts of overlapping crises. She is passionate about leveraging data and technology for a more peaceful and just world. Her work extends from the independent evaluation offices of UNDP and OIOS in New York to monitoring roles in humanitarian crisis and conflict affected regions for WHO and MINUSMA. From 2019 to 2022 she co-convened the UNEG Working Group on Methods, a role indicative of her dedication to pilot new methods and foster innovation and exchange in the community of practice.

Heidi Larson: Heidi J. Larson, PhD, is Professor of Anthropology, Risk and Decision Science and is the Founding Director of the Vaccine Confidence Project at the London School of Hygiene & Tropical Medicine. She is also Clinical Professor of Health Metrics Sciences, University of Washington, Seattle, USA, and Guest Professor at the University of Antwerp, Belgium. Dr. Larson previously headed Global Immunisation Communication at UNICEF, chaired GAVI's Advocacy Task Force, and served on the WHO SAGE Working Group on vaccine hesitancy. She served on the FDA Medical Countermeasure (MCM) Emergency Communication Expert Working Group, and is currently Principal Investigator for a global study on acceptance of vaccination during pregnancy; an EU-funded (EBODAC) project on the deployment, acceptance and compliance of an Ebola vaccine trial in Sierra Leone; and a global study on Public Sentiments and Emotions Around Current and Potential Measures to Contain and Treat COVID-19.

Christophe Legrand: Christophe Legrand is an Evaluation Specialist at the United Nations Capital Development Fund (UNCDF). He has more than 10 years of experience at Country, Regional and Headquarters level and a strong background in evaluation methodologies and evaluation management, as well as ample experience with implementing monitoring strategies to support project, programme

and organizational learning on wide range of topics. Knowledgeable of bilateral and multilateral donor requirements, he thrives in challenging environments, is resilient and enjoys working in a team. He collaborated with the methods working group in his capacity as coordinator, Evaluating Policy Influence Working Group.

Sergiu Tomsa is a Social and Behaviour Change Specialist with UNICEF Regional Office for Europe and Central Asia, providing technical advice, guidance and support to UNICEF country offices and national stakeholders in generating social and behaviour evidence, designing and testing behavioural insights solutions, developing and implementing behaviour and social change strategies in the areas of immunization, early childhood development, social inclusion, prevention of violence against children, parenting and other. Before joining UNICEF Regional Office, Sergiu has worked with UNICEF Office in Kosovo (UNSCR 1244) and UNICEF in Moldova. Prior to UNICEF, Sergiu has worked with the Organization for Security and Cooperation in Europe (OSCE), and with child's rights non-government organizations in Moldova. He has a master degree in Social Psychology.

Deborah McWhinney: Deborah McWhinney is a results-driven professional with more than twenty-five years of leadership experience in strategic policy and programme planning, analysis, and evaluation with bilateral and multilateral agencies, international NGOs and the private sector. For over ten years, she has applied an innovative and effective management style to oversee a range of complex, global evaluations, programmes, and initiatives in diverse settings and sectors. She is currently working as an Evaluation Advisor at UNFPA after 6.5 years as a Senior Evaluator and Head of the Global and Synthesis Unit at the World Food Programme's Office of Evaluation. Prior to joining WFP, she worked for the Development Evaluation Unit at Global Affairs Canada, coordinated assessments of multilateral organisations on behalf of MOPAN and managed the International Program for Development Evaluation (IPDET). She worked for UNICEF for 10 years and managed programmes for CARE International. She holds a Master's degree from York University and a Diploma in Public Policy and Program Evaluation.

Veridiana Mansour Mendes: Veridiana Mansour Mendes has over 10 years' experience in evaluation, policy analysis and social research. She is currently working as Evaluation Officer at WIPO, where she conducts thematic and strategic evaluations. Prior to joining WIPO, she worked with FAO, governments, private companies and civil society organizations in a variety of themes and policy sectors. She is particularly interested in topics surrounding policy coherence and evidence-based policymaking, and has been co-coordinating the UNEG sub-working group on evaluating policy influence since 2018. She holds a bachelor's degree in Law from the Pontifical Catholic University of Sao Paulo and a master's degree in Public Policy from the University of Edinburgh.

Roger Miranda: Roger Miranda is Evaluation Officer at the Food and Agriculture Organization. With over twenty-five years of experience in programme evaluation, his interests include making evaluation accessible to non-experts, having authored "Eva the Evaluator" as a means to bring the topic to a wider audience. He is a member of the Board of the Center for Theory of Change.

Linda Morra Imas: Dr. Linda Morra Imas is a founder and first President of the new International Evaluation Academy. She is also co-founder of the International Program of Development Evaluation Training, IPDET, and served as its co-director and chief instructor for 17 years. She was a Chief Evaluation Officer and Evaluation Capacity Building Adviser for the World Bank Group, working both in its private and public sectors, and a senior director at the U.S. Government Accountability Office. Now an independent consultant, she advises on M&E and provides training worldwide, building on 30+ years of experience. Among other publications, she co-authored the text on development evaluation, *The Road to Results: Designing and Conducting Effective Development Evaluations*, now translated into eight

languages. She is an honorary member of the International Development Evaluation Association, IDEAS, and is also known for her work on development evaluation competencies and professionalization.

Nora Murphy Johnson: Dr Nora Murphy Johnson, with more than 10 years of experience in research, evaluation, and systems change, has conducted numerous evaluations, published articles, and delivered workshops on the use of developmental evaluation and other systems change approaches. Of note, she was the recipient of the Michael Scriven Dissertation Award for Outstanding Contribution to Evaluation Theory, Method, or Practice based on her work with using developmental evaluation for social justice. More information about Nora's work and organization can be found [here](#).

Sandy Oliver: Dr Sandy Oliver is Professor of Public Policy at UCL Institute of Education. For thirty years her interests have focused on the interaction between researchers and people making decisions in their professional and personal lives, largely through the conduct of systematic reviews. She is a member of the Board of the Campbell Collaboration and Cochrane editor with their Consumers and Communication Review Group. She works with the UK Department for International Development to build capacity in systematic reviewing in developing countries.

Elisabetta Pegurri: Elisabetta Pegurri is a Senior Evaluation Adviser at UNAIDS. She has over 20 years of work in public health, including in African countries such as Mali, Rwanda and Ethiopia. She has a background in health economics and psychology. She is an evaluation and M&E expert with strong collaboration and partnership skills. She provided strategic leadership for the creation of the first-ever independent evaluation function at UNAIDS. She has managed and co-conducted multiple complex evaluations and led joint evaluations. She is proficient in quantitative data analysis and statistical methods, and various evaluation approaches, including theories of change, qualitative interviews, and focus groups, as well as surveys – all with a particular focus on effective engagement of rights holders and vulnerable groups.

Saltanat Rasulova: Dr Saltanat Rasulova is a Regional Evaluation Specialist at UNICEF, based in Kazakhstan. Previously, Saltanat held positions at Oxford Policy Management, University of Bath, and University of Oxford. Saltanat received a PhD in Social Policy from Oxford University.

Patricia Rogers: Patricia Rogers is Chief Executive Officer of Better Evaluation, an international collaboration, NGO and registered charity with a mission to improve monitoring and evaluation by creating, sharing and supporting the use of knowledge about evaluation methods, processes and approaches. She was formerly Professor of Public Sector Evaluation at RMIT University and is now visiting professor at the University of Witwatersrand, South Africa. She is an experienced evaluator who has worked for more than 30 years in a wide range of sectors, countries and organisations, including national and sub-national government, UN agencies, development banks, NGOs and philanthropic organisations. Her PhD developed a framework for evaluating approaches to program evaluation, and she remains committed to improving evaluation to maximise its benefits and manage its risks. She is a Fellow of the Australian Evaluation Society and a recipient of the American Evaluation Association of evaluation, the AES Best Evaluation Study, the American Evaluation Association's Alva and Gunnar Myrdal Practice Award.

Zoe Sutherland: Zoë Sutherland is a Senior Consultant at ITAD and has over 10 years' experience in international development, with specialisms in developing outcome monitoring systems, evaluating large and complex portfolios and providing strategic support to impact measurement. Prior to joining Itad, Zoë worked as a Senior Impact Advisor in Business Programs at Save the Children, designing results frameworks, monitoring systems and impact dashboards for businesses on the corporate benefits and social impact of NGO-corporate partnerships.

Carlos Tarazona: Carlos Tarazona is a development specialist with over twenty years of experience in the design, management and evaluation of agricultural and rural development programs. He was co-chair of the UN Evaluation Group strategy task force (in charge of developing the UNEG Strategy 2020-24), and is currently a Senior Evaluation Officer in FAO and co-coordinator of the UNEG interest group on methods.

Tina Tordjman-Nebe: Dr. Tina Tordjman-Nebe is Senior Evaluation Specialist at UNDP's Independent Evaluation Office (IEO) in New York, where she conducts (and acts as senior advisor to) corporate/thematic evaluations and independent country programme evaluations. She leads efforts to boost organizational learning via IEO's newly established Synthesis & Lessons unit, including through the SDG Synthesis Coalition to inform deliberations on the post-SDG development agenda and the IEO Reflections series, offering lessons based on rapid evidence assessment of past evaluations. She has a longstanding interest in methodological rigour and appropriateness and has co-coordinated the UNEG working group on methods since its inception in 2019.

Pietro Tornese: Pietro Tornese is an Economic Affairs Officer with the UN Economic and Social Commission for Western Asia (UN ESCWA), where he supports policy making in the Arab region across the humanitarian, development and peace nexus. Prior to joining UN ESCWA, Pietro served as Evaluation Analyst at the UN Capital Development Fund in New York. He has also worked for Ernst & Young, Open Evidence, and the Italian Agency for Development Cooperation in Latin America and the Caribbean. As an independent consultant, Pietro has co-authored a number of research and evaluation reports for the European Commission.

Vijayalakshmi Vadivelu: Vijayalakshmi Vadivelu is an Evaluation Advisor at the Independent Evaluation Office of UNDP. She has managed and carried out a number of evaluations of development and crisis related programmes at the global and country levels. Prior to joining UNDP she carried out applied research on governance issues and gender equality in development at the Institute for Social and Economic Change, Bangalore, India. Vijaya holds a Ph.D. in Sociology.

Henri van den Idsert: Henri van den Idsert is a Senior Evaluation Officer at UNHCR who oversees a broad range of evaluations under his portfolio, including country strategy evaluations, livelihoods, inter-agency humanitarian evaluations, and emergency preparedness and response. Henri has over twelve years of experience in monitoring and evaluation, his experience extends across the humanitarian-development nexus, having worked for UNDP in South Sudan, the Dutch Ministry of Foreign Affairs in the Hague, NGOs in the humanitarian and sexual reproductive health and rights sector, as well as public policy and development consultancies in Somalia and Kenya.

Pankaj Kumar Verma: Pankaj Verma is an Evaluation Officer at the United Nations Office of Internal Oversight Services (OIOS). He has 11 years of experience leading evaluation studies, programme implementation, data collection and analysis and policy advocacy across a portfolio of projects on the issues of women health, energy, public employment, education, health insurance, and extraction industries, including impact evaluations and large-scale randomized control trials (RCTs). Prior to joining OIOS, he worked as an Economist & Coordinator at Social Observatory at the World Bank in India, as a Country Economist with International Growth Centre, at J-PAL on various large-scale randomized control trials, and with the World Bank in Cambodia to develop a trade profile for the country.