

EVALUATIONS
CONNECTIONS

Jos Vaessen Guest editor. Ole Winckler Andersen Associate Editor. Tereza Zvolská Editorial Assistant.

EES

EUROPEAN
EVALUATION
SOCIETY

APRIL 2016

ARTICLES – NEWS – EVENTS

CONTENT

Presidential message	1
Evaluation practices in the United Nations An editorial	2
The United Nations Evaluation Group: Helping the United Nations build a better world	4
Special evaluation challenges facing UN agencies in the age of the Sustainable Development Goals, complexity and Big Data	5
An analysis of the evaluation function in the United Nations System	6
Evaluation use in the United Nations System	8
Evaluating IFAD's support to rural development in fragile states and situations	12
Evaluating standard-setting work in UNESCO: Culture conventions	14
Evaluation and the global environment: Challenges and opportunities	16
The authors	18
Guidance to contributors	20

PRESIDENTIAL MESSAGE

Riitta Oksanen

Dear EES members and friends,

Inspired by the focus of this Connections issue on evaluation in the United Nations system this letter seeks to raise two topics that may merit further discussion and follow-up action.

The first concentrates on the potential value added by the United Nations Evaluation Group (UNEG) not only within the UN system but more broadly in United Nations member countries and particularly in Europe. Specifically, I wonder whether evaluation societies in Europe (including EES) could do more to promote the values and ideals of the United Nations, including a human-rights-based approach in evaluation.

My day-to-day job at the Finnish Ministry for Foreign Affairs is to ensure that we have a development evaluation system that ensures independent, credible and useful evaluations. For a relatively small evaluation unit like ours, international evaluation norms and standards have been the most important resource for the development of our evaluation function. We have over many years systematically benchmarked our system against the OECD-DAC norms and standards. This has given us valuable guidance on where our

challenges are, and where to focus our own capacity building efforts.

Finland's development policy has traditionally emphasized the promotion of human rights, particularly in the area of promoting gender equality and women's rights. In this area we have made full use of the tools developed by UNEG. The guidance on "Integrating Human Rights and Gender Equality in Evaluation" has been helpful in providing evaluation practitioners with hands-on and implementable ideas on how to commission and conduct evaluations that can make a difference.

Shouldn't this approach be adopted by all European countries? Gender equality and social equity need strengthening in Europe as in other parts of the world and closer cooperation between the UN evaluation system and European evaluation societies may help to unleash the transformative power of evaluation throughout the European continent. What mechanisms and initiatives would be needed to trigger such a development as European countries pursue the implementation of the universal Sustainable Development Goals (SDGs)?

My second topic is best introduced by asking a simple question to UN system evaluation

colleagues: why is it so hard for UN agencies to “Work as One”? Would the new impetus of the SDGs and the call for more intensive evaluation capacity development efforts at the country and thematic levels provide an opportunity for a new start? We all know the “One UN” slogan that has served as a major driver for UN reform. In the field of evaluation UNEG has a strong and important coordinating role when it comes to norms

and standards, and to monitoring UN wide progress. A lot has been achieved. The most difficult step, however, has not been taken, i.e. the giant leap that would be needed to implement and evaluate together.

Established institutions, power structures, and practices are not easy to change. But many UN evaluation offices are becoming active in support of evaluation capacity de-

velopment efforts at country level. There are also innovative initiatives that seek to amplify the voices of all stakeholders in the public sector, civil society and parliaments.

Wouldn't tackling the country-led evaluation challenges raised by the SDGs be one way of inducing greater momentum behind the “UN as One” vision? ■

EVALUATION PRACTICES IN THE UNITED NATIONS

AN EDITORIAL

Jos Vaessen

In 1951, Professor Walter Sharp published an article about the role of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in a post-World War II context of fragile international relationships and growing tensions between the East and the West (Sharp, 1951). A number of his arguments regarding UNESCO's mission as a multilateral institution – e.g. nurturing an intercultural understanding among Member States – are as relevant today as they were then. Sharp's evaluative assessment of UNESCO can be considered as one of the first ‘evaluations’ of a United Nations (UN) organization at the corporate level¹.

In the late seventies the Joint Inspection Unit (JIU) carried out a succinct assessment of the state of evaluation in the UN (JIU, 1977). Evaluation in the UN in that period could best be characterized as an ‘infant industry’. Evaluations were mostly implemented in an *ad hoc* manner without much guidance or support in terms of organizational structure, procedures or staff capacities. Most (of the then existing) UN organizations started experimenting with evaluations in the 1960–70s. Some even had a rudimentary evaluation framework in place from the very start of their operations (e.g.

UNIDO). UNDP and WHO and a few other organizations were frontrunners in terms of developing dedicated evaluation structures and processes, which were primarily focused on the project level.

Gradually, the UN moved forward toward institutionalization of the evaluation function. In 1984, the Inter-Agency Working Group on Evaluation was established. By the turn of the century this led to the establishment of the United Nations Evaluation Group (UNEG), a voluntary network bringing together the evaluation functions of UN Secretariat departments, funds, programs, specialized agencies, regional commissions and others. The establishment of UNEG provided a big boost to the formalization and harmonization of evaluation functions and practices across the UN system. Among other things, UNEG developed norms and standards for evaluation in the UN; guidance on evaluation policies, processes, methodologies and practices; and a platform for information exchange among UNEG members.

In the first contribution to this Special Issue, *Marco Segone* discusses the role of UNEG at the dawn of the 2030 Agenda for Sustainable Development, presenting a case for more

cooperation among the international networks and partnerships for evaluation.

The subsequent three contributions discuss cross-cutting evaluation challenges in the UN. *Michael Bamberger*, acknowledging some of the specificities of evaluation in the UN system, identifies five sets of challenges that UN evaluation units should address. One of these concerns evaluation in the context of the recently adopted Sustainable Development Goals. The fact that many evaluations are confined to a short-term and project-bound perspective is unhelpful. Among other things, there is a need for evaluations that adopt a long-term and systemic perspective to illuminate the contributions of UN policy interventions to the attainment of the SDG targets.

Next, *Sukai Prom-Jackson and George Bartisiotas* of the JIU present an analysis of the maturity of the evaluation function in the UN system. In stark contrast with the analysis presented in the 1977 JIU report, most UN organizations now have a firmly institutionalized central evaluation function. Yet many challenges remain. For example, the majority of evaluations conducted in the UN system are decentralized evaluations which are not

directly under the control and management of the central evaluation function. While corporate evaluations (which are under the control of the latter) are embedded in clear frameworks of evaluation planning, design, implementation and follow-up and supported by qualified staff, more often than not this is not the case for decentralized evaluations.

A final cross-cutting UN perspective comes from *Robert Stryk*. He discusses recent work conducted in the framework of UNEG on evaluation use. This joint exercise, with the participation of several UNEG members, identified six (expected) patterns of evaluation use in the UN. The challenge lies in identifying under what conditions it is more or less likely that these patterns unfold as expected. Readers may wish to read this contribution and the previous one in tandem with *Indran Naidoo's* interesting perspective on independence of evaluation in the UN, which was published in the February issue of *Connections*.

The second part of the Special Issue comprises three contributions presenting UN agency-specific experiences of evaluation. *Oscar Garcia*, *Ashwani Muthoo* and *Fabrizio Felloni* discuss the approach and findings of a recent evaluation of the International Fund for Agricultural Development's (IFAD) support to rural development in fragile states and situations. Particularly interesting are the delineation challenges faced by the evaluation, as the evaluand covers a range of

interventions in different types of situations of fragility. The evaluation is based on a comprehensive mixed methods approach that combines portfolio-level evaluative analysis with case studies.

A set of evaluations of UNESCO's standard-setting work in culture is the subject of *Barbara Torggler's* contribution. The UN plays an important role in setting international standards for different policy fields. Yet, in contrast to other areas of work, the evaluation community has struggled with developing adequate approaches. This contribution helpfully discusses some of the complex issues that evaluations need to address in the context of UNESCO's culture conventions. The contribution also provides some pertinent points on evaluation use in an international stakeholder community that is not so familiar with evaluation.

Finally, *Juha Uitto* discusses the challenges and opportunities for evaluating interventions aimed at influencing the global environment (e.g. biodiversity, climate change, international waters) from the perspective of the Global Environment Facility (GEF). Environmental change processes and the human agency-environment nexus are inherently complex, requiring methodological designs that go beyond conventional approaches. The contribution clearly demonstrates the need for methodological innovation, using the example of a recent GEF impact evaluation. Employing a multi-level, multi-method

approach – combining such methods as remote sensing, statistical counterfactual analysis, qualitative case studies, and several other methods – the evaluation was able to capture the contribution of GEF support to biodiversity conservation.

This Special Issue of *Connections* is not intended to present the reader with a comprehensive picture of evaluation in the UN. For example, it lacks discussion of evaluation in the context of the UN's political, humanitarian and peace-building work. Yet, the cross-cutting and agency-specific perspectives on evaluation that this issue provides should inform readers about the diversity and vitality of evaluation practices in the UN. Since *Sharp (1951)*, evaluation in the UN has come a long way. However, as shown in this Special Issue, evaluation as a function and a field of practice needs to constantly adapt and evolve in line with the challenges presented by the fast-paced and increasingly interconnected world that we live in.

References

- JIU (1977). *Report on evaluation in the United Nations system*. Geneva: Joint Inspection Unit.
- Sharp, W.R. (1951). The Rôle of Unesco: A Critical Evaluation. *Proceedings of the Academy of Political Science*, 24(2), 101–114.

1 At about the same time, the UN Secretariat in New York professed a need to evaluate the UN's technical assistance work, but evaluation as an institutionalized function in the United Nations did not exist at that time.

THE UNITED NATIONS EVALUATION GROUP: HELPING THE UNITED NATIONS BUILD A BETTER WORLD

Marco Segone

2015 was a historic year. It was officially designated as the International Year of Evaluation (EvalYear). This prompted a global movement spanning from the UN Secretary General Ban Ki-Moon to emerging evaluators. Moreover, the new 2030 Agenda for Sustainable Development was launched by 192 Heads of State at the UN Sustainable Development Summit in New York in September 2015.

The new 2030 Agenda encapsulates the Sustainable Development Goals. It constitutes a universal and transformative agenda for the next fifteen years that is relevant not only for developing countries but also for developed nations and regions, including Europe. The new global goals focus on people, the planet, prosperity, peace and partnership. They highlight the importance of gender equality, social justice and equity, and promise that no one, especially the most vulnerable, will be left behind.

How then will the world realize the goals of a people-centered agenda for sustainable and equitable development? What do we need to know? What do we need to do differently and better? The world's nations and their citizens need to own these questions and their answers. This is a key element in a new paradigm of accountability, national ownership, partnerships, and transparency. People have a right to know about and participate in the decisions and actions which affect them. "Meeting people's right to evidence" emerges as a rallying cry for policy makers and citizens alike.

Evaluation can play a transformational role here. How do we know if policies and programs are effective and, most importantly, are reaching the most disadvantaged communities, families and individuals: women and men, girls and boys? How do we know if public money is being used efficiently and wisely? What is working and not working, and why? Evaluation can generate the evidence to answer these and many other ques-

tions of public interest. But it is important that governments and citizens have access to such evidence, and use it to inform public debates and decisions and to resist harmful or wasteful actions and interventions. To help provide such evidence and promote its use, most UN entities have an evaluation office. Yet, working alone, these units cannot realize the full potential of evaluation and evaluation evidence.

This is where UNEG – the United Nations Evaluation Group – comes in. As a professional network, linking evaluation offices across 47 different UN entities, it sets quality standards, provides professional support and amplifies the voices of UN evaluators in their various organizations (UNEG, 2016). This in turn enables evaluation offices in the different United Nations entities to provide strategic and meaningful contributions to the global community, the United Nations system, and their own organizations. In this way UNEG helps the United Nations to sharpen and strengthen its relevance, efficiency and effectiveness through appropriate and well-informed decisions.

For these reasons, UNEG works for a strong United Nations system-wide mechanism which will provide evidence of what works and what does not work in system-wide initiatives, including UN reform and system-wide action on gender equality. UNEG supports each member, no matter its size, resources and capacities, to ensure that UN evaluation functions provide credible and reliable evaluation evidence to inform and strengthen the UN, in both development and humanitarian contexts; in stable as well as fragile countries; and, ultimately, for all the world's people.

Of course UNEG cannot do this alone. The same goes for the UN that UNEG serves, as well as individual governments. As the UN Secretary General stated in his "The Road to Dignity by 2030" report: "The new agenda must become part of the contract between

people, including civil society and responsible business, and their governments – national and local" (UN, 2014). The emerging paradigm requires global alliances and partnerships to build tomorrow's world.

This is why UNEG is engaging with a wide range of stakeholders, united under the banner of EvalPartners and EvalGender+, a global partnership to strengthen national capacities for equity-focused and gender-responsive evaluation. EvalPartners and EvalGender+ bring together Voluntary Organizations for Professional Evaluation (including the EES), parliamentarians, policy makers, United Nations agencies, multilateral banks, private foundations and civil society organizations.

The goal is to strengthen capacities in countries to generate good quality evidence to inform their own national development strategies. And UNEG is a key member, committed to make a difference – as already demonstrated by its successful advocacy of including evaluation in the 2030 Agenda approved by the UN General Assembly.

In a nutshell, UNEG serves the UN in building a world that works: that is, a world free of poverty, discrimination and gender inequalities; a world of peace, social justice and respect for the environment. Within the UN system and beyond, UNEG in collaboration with its partners will continue to reaffirm and strengthen the role of evaluation toward achieving these ambitious goals.

References

UN (2014). *The road to dignity by 2030: ending poverty, transforming all lives and protecting the planet*. Retrieved 4 February 2016, from http://www.un.org/ga/search/view_doc.asp?symbol=A/69/700&Lang=E

UNEG (2016). *United Nations Evaluation Group*. Retrieved 4 February 2016, from <http://unevaluation.org>

SPECIAL EVALUATION CHALLENGES FACING UN AGENCIES IN THE AGE OF THE SUSTAINABLE DEVELOPMENT GOALS, COMPLEXITY AND BIG DATA

Michael Bamberger

In addition to the wide range of methodological and real-world economic, political and organizational challenges that all evaluation departments face (Bamberger et al., 2012), UN evaluation offices must address challenges resulting from the broad nature of their mandates and the fact that they are asked to address the most severe problems facing the developing world. I identify five big challenges. Unfortunately, space does not permit a review of the significant advances that are being made in most of these areas. In particular, the widespread endorsement of the Sustainable Development Goals (SDGs) has already stimulated a great deal of exciting and important work in areas such as Big Data, equity, and gender equality (to name only three) as well as a rethinking of broader evaluation strategies.

A first set of challenges concerns the development and implementation of an evaluation framework for the SDGs. The evaluation of the SDGs introduces unique challenges. The focus on sustainability presents two new dimensions. Sustainability requires that the evaluation must cover a much longer period of time as many programs have an expected lifetime of 5–10 years or more, and sustainability should be assessed over this whole period. Also, while conventional evaluations are based on pre-test – post-test designs using data that measure events that have already taken place, addressing sustainability requires looking into the future and estimate the likelihood that a program will continue to operate and to deliver benefits and services. Currently, many evaluations are funded under a project modality and when the project ends evaluation funding is also terminated. This provides a challenge for collecting ongoing data to assess sustainability over time.

Sustainability also requires multi-sectoral analysis as programs do not operate in a vacuum. For example, the sustainability of health programs depends on what is happening in education, food and nutrition,

women's rights and transportation – among others. A challenge for SDG evaluations is to find time to distill the lessons from the huge number of evaluations relating to the Millennium Development Goals in time to be able to build on these lessons, both successes and failures, for the SDG evaluations. Another major challenge is to develop an integrated evaluation design, data collection and analysis for the 17 SDGs and developing standardized indicators that permit cross-country comparisons while capturing the unique features of each country.

A second set of challenges concerns the need to respond to the increasing demands to address equity, social justice and gender equality issues. While significant progress has been made on all of these issues, none have been adequately mainstreamed into evaluation practice. With respect to gender, almost all development agencies now have gender policies and gender action plans and most have guidelines on how to incorporate gender into their evaluations. However, often attention to gender tends to decline as the project cycle progresses. So while gender is frequently addressed at the earlier diagnostic stage, and perhaps in project design, attention is weaker in project implementation and often even weaker in project evaluation. While most agencies recognize this, even if they do not fully address differences between women and men, very few agencies have made much progress on the more complex issues concerning the lesbian, gay, bisexual and transgender communities. Equity has received increased attention in recent years, but many agencies still focus on aggregate statistics concerning, for example, the proportion of children not attending school or the proportion of the population below the poverty line; but it has proved more difficult to recognize that the gap between the rich and poor may be increasing (even while aggregate statistics improve) or to distinguish between poverty and destitution. So often the situation of the bottom, say 20 per cent, is not adequately addressed (Bamberger & Segone, 2011).

A third challenge concerns the inability of many evaluations to adequately identify and address unintended outcomes. Work on gender provides many examples where programs to promote women's economic and social empowerment, for example through providing access to mobile phones or small business development support, can increase domestic violence from men who feel threatened. While these negative outcomes are well understood in the gender community, they are often not captured in many evaluations, particularly the large number of rapid evaluations where consultants only have time to meet with women who have benefited from the programs, and often do not have time or the incentive to dig deeper into the darker side.

Most evaluations are designed to assess the extent to which programs have achieved their *intended* outcomes and many widely used evaluation designs are unable to identify even very serious unintended outcomes. This criticism also applies to many of the supposedly more “rigorous” evaluations such as randomized control trials which are designed to focus narrowly on statistically significant differences between treatment and control groups on specific program (intended) outcomes (not unintended outcomes). Many results-based monitoring and evaluation systems are also subject to this criticism. Unfortunately, experience suggests that many agencies are content with receiving positive evaluations of their programs and some even discourage evaluators from seeking to identify or report unintended, and frequently negative, outcomes.

Fourth, evaluation offices are only just beginning to address the fact that the increasing complexity of development interventions poses significant challenges to conventional evaluation approaches. Most agencies have not yet developed frameworks either to clearly define the extent to which their programs and the contexts in which they operate can be considered complex, and even less to the devel-

opment of complexity-responsive evaluation designs (see Bamberger et al., 2016).

Finally, evaluators are only just beginning to recognize the substantial implications of the world of Big Data for the evaluation profession. While the importance of Big Data has grown dramatically over the past five years (UN Global Pulse, 2012) much less attention has been given to integrating Big Data into program monitoring and evaluation (Letouzé et al., 2016). A major challenge for the SDG evaluation strategies will be to recognize and harness the power of Big Data. One example concerns the need to develop dynamic integrated data platforms to permit comparability across different sectors.

Conventional approaches to data platforms involve working with well-defined indicators based on surveys or government statistics that are well-understood and relatively stable. However, the integration of Big Data involves a fundamentally different approach

where large numbers of rapidly changing indicators, whose quality cannot be controlled, but which have tremendous potential value, must be integrated into the data platforms. For example, how to integrate indicators from Twitter on food security which change daily and whose quality and representativity is difficult to assess, with carefully designed and implemented national food surveys which are very detailed and statistically representative, but where the information is many months out of date by the time it is published?

Addressing these challenges requires a concerted effort by UN evaluation offices in tandem with the broader development (evaluation) community.

References

UN Global Pulse (2012). *Big Data for Development: Opportunities & Challenges*. Retrieved

8 January 2016, from <http://www.unglobalpulse.org/projects/BigDataforDevelopment>

Bamberger, M., Rugh, R. & Mabry, L. (2012). *Realworld Evaluation: Working under budget, time, data and political constraints*. Thousand Oaks: SAGE.

Bamberger, M. & Segone, M. (2011). *How to design and manage equity-focused evaluations*. New York: UNICEF.

Bamberger, M., Vaessen, J. & Raimondo, E. (eds.) (2016). *Dealing with complexity in development evaluation: a practical approach*. Thousand Oaks: SAGE.

Letouzé, E., Areias, A. & Jackson, S. (2016). The evaluation of complex development interventions in the age of big data, in: Bamberger, M. Vaessen, J. & Raimondo, E. (eds.) op.cit.

AN ANALYSIS OF THE EVALUATION FUNCTION IN THE UNITED NATIONS SYSTEM

Sukai Prom-Jackson and George Bartsiotas

The United Nations is at a crossroads and should consider major transformations to strengthen its leadership role in global governance, enhance transparency and accountability for results, and demonstrate its added value. Evaluation is a critical oversight mechanism that would help the UN to undertake these necessary transformations. Building on its distinguishing feature of critical inquiry, it could play a key role in helping the UN system account for its results (intended or unintended) and understand whether it is doing the right things, doing things right, as well as the sustainability of its activities and results. Moreover, evaluation would help the UN understand where its comparative added value and strategic significance lie and how to get there.

In 2014, the Joint Inspection Unit (JIU) conducted a study to assess the extent to which the evaluation function in the UN system is able to perform these roles. It also looked

at the factors affecting the capacity and performance of the evaluation function. Finally, it examined the readiness and adaptability of the function to respond to emerging changes and challenges, including the implications for evaluation of the Post-2015 development agenda.

The study sought to contribute to ongoing efforts across the system to strengthen the capacity of the evaluation function in meeting professional standards, and more generally in enhancing the value of the United Nations system in the world. The report fed into the 2015 International Year of Evaluation, the implementation of the UNEG strategy on evaluation, as well as the imperatives of the post-2015 development agenda.

The study covered 28 UN system organizations: Funds and Programs, Specialized Agencies, special technical entities, and entities from the UN Secretariat. The methodology

relied principally on a self-assessment exercise conducted by each UN system organization, which was validated and triangulated with additional evidence from an elaborate desk study and semi-structured interviews with representatives from the different UN system organizations.

More specifically, the approach used in the study relied upon: (i) a validated standardized assessment framework articulated to the United Nations Evaluation Group norms and standards for evaluation and for institutional development; and (ii) an assessment of the maturity of the evaluation function. Regarding the latter, the study distinguished between five levels of maturity, whereby each maturity level corresponds to an aggregate rating on a number of criteria, supported by detailed semantic scales (for the detailed approach see JIU, 2014). Finally, the methodological approach relied on the principles of holism (the whole is more than the sum of

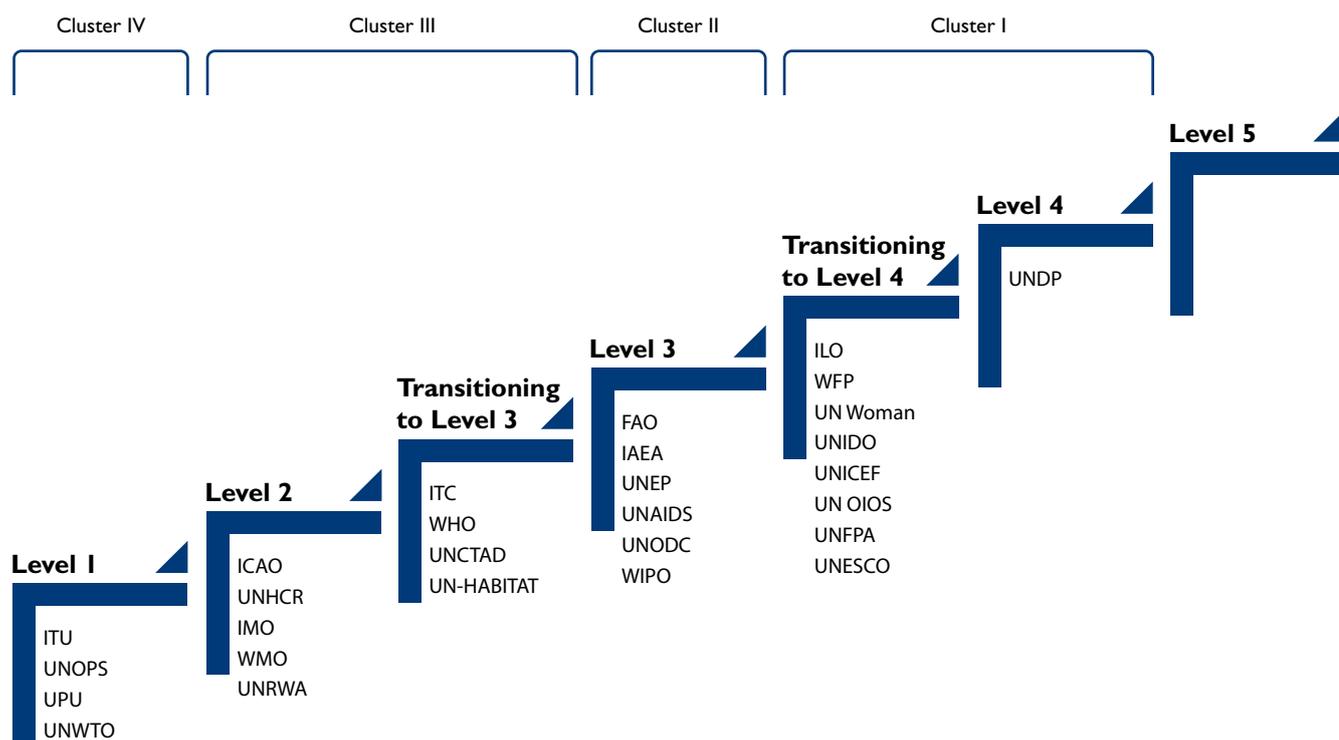


Figure 1. Overall level of maturity of the central evaluation function for 28 JIU participating organizations.

	Small		Medium		Large	
	Organization	Grade	Organization	Grade	Organization	Grade
Stand alone	UNODC	5.2	ILO	6.7	WFP	6.7
	UNIDO	6.4	UNEP	5.8	UNDP	7.1
	UN Women	6.4	UNFPA	6	UNICEF	6.3
	UN-Habitat	4.2			FAO	5.9
Co-located/ with management	ITC	4.9	UNAIDS	5.3	UNHCR	3.8
	UNCTAD	4.4				
Co-located with oversight/ audit	IMO	3.7	UNESCO	6.1	UN-OIOS	6.2
	WMO	3.5	UNRWA	3.4	WHO	4.7
	ICAO	3.9	WIPO	5.2		
			IAEA	5.9		

3.0–4.9 = Below average and low
 5.0–6.0 = Average
 6.1–8.0 = Above average and high

Table 1. Level of maturity of the evaluation function by size of the organization, overall annual budget and the location of the central evaluation function.

the parts), and non-summativity (no system is stronger than its weakest link).

The report focuses mainly on the central evaluation functions of the 28 entities, which support corporate-level policy and strategic decision-making. It also provides a rapid review of decentralized evaluation functions,

which are generally embedded in program and operational units throughout the United Nations system, supporting line management decision-making.

The report found that central evaluation functions in the 28 organizations have significantly evolved over time, generally becoming

more efficient and providing better quality evaluation services. However, the level of commitment to evaluation across the United Nations system is not commensurate with the growing demand and importance of the function. Moreover, very few organizations have clearly defined institutional frameworks for decentralized evaluations. Figure 1 re-

ports the aggregate results of the comparative assessment of the level of maturity of the central evaluation functions of the 28 organizations covered by the study.

In most cases the level of maturity is influenced by the size of the organization, the resources committed to evaluation, and the structural location of the function. The latter dimension is illustrated in Table 1. As shown in the table, there is a weak association between maturity and size. In addition, there is a weak association between the location of the central evaluation function and the level of maturity.

The report presents a number of recommendations to improve the evaluation function in the United Nations system. They include, inter alia:

- the need for organizations to deal with systemic constraints associated with the function being under-resourced and over-

stretched and, therefore, unprepared to respond to emerging demands;

- the need for organizations to be more strategic in positioning the function to play a greater role as an agent of change and in balancing priorities between accountability and the development of a learning organization; and
- the need for organizations to think in terms of the UN system as a whole and on how they can work in an integrated fashion.

The evaluation also calls for more openness to restructuring the evaluation function to support change and transformation in today's world, including:

- enhanced strategic linkages between the central evaluation function and decentralized evaluation;
- more joint evaluations, more system-wide evaluations of operational activities for

development, and more common mechanisms for quality assurance;

- enhanced linkages with national evaluation systems and support for national evaluation capacity development.

Finally, with respect to decentralized evaluation functions, the report calls for a dedicated effort, a clear institutional framework, and more resources to further clarify and improve the role and contribution of decentralized evaluations to learning and accountability in the UN system.

References

JIU (2014). *Analysis of the Evaluation Function in the United Nations System*. Retrieved 11 January 2016, from https://www.unjiu.org/en/reports-notes/JIU%20Products/JIU_REP_2014_6_English.pdf

EVALUATION USE IN THE UNITED NATIONS SYSTEM

Robert Stryk

All evaluations are undertaken to be used. Indeed, Patton (1997, in Herbert 2014, p.389) claims that evaluations that are not used, regardless of their quality, tend to be considered failures. The United Nations Evaluation Group (UNEG) in its 2014 - 2019 Strategy recognized the importance of use of evaluations and defined the second strategic objective of the Strategy as “*United Nations entities and partners use evaluation in support of accountability and program learning*” (UNEG, 2014).

To better understand how United Nations agencies use evaluation and to identify the factors that support and hinder evaluation use, a working group was established to identify key mechanisms that are expected to enhance the use of evaluations in the United Nations System. The group gathered data from four different sources: (a) a literature review of the most relevant academic

and non-academic sources on evaluation use; (b) an online survey of United Nations evaluation users and practitioners, as well as external evaluation practitioners; (c) semi-structured interviews with United Nations evaluation users and practitioners; and (d) case studies of instances from the UN system where evaluation was useful and used.

The work confirmed the factors that enhance use as mentioned in the literature, as well as confirming that the level of use of evaluation needs to be improved. Results of the survey conducted showed that only a quarter of respondents rated use as “high” (defined as ‘systematic use for decision-making, new project cycles, corrective actions’), another quarter of respondents rated use as “low” (defined as ‘infrequent use’), with the remaining respondents rating use of evaluation in the United Nations System as “medium” (defined as ‘periodic use’).

The work is summarized in six key messages that were developed and then tested through presentations to peers (fellow evaluators in the UN system). The messages, as well as the underlying mechanisms explaining evaluation use, are intended to provoke discussion and to be modified as the understanding of the mechanisms improves over time.

Message 1: Users and stakeholders should be involved in and consulted throughout the evaluation process

Evaluators have long known about the importance of involving users and (other) stakeholders in the evaluation process. This has been a focus of the evaluation literature for the past decades, and the message seems to have been understood. There are several potential benefits from involving stakehold-

ers in all aspects of the evaluation process: (1) their involvement and consultation means that they develop more ownership over the findings; (2) the consultation process improves the quality of the recommendations; and (3) consultation and engagement increase process use. The main expectations about intended causal processes underlying this message are depicted below.

Message 2: The support of senior decision makers is key, and so is their commitment to the implementation of the recommendations

In all organizations, support from the top helps drive change. The evaluation world is no different. Evaluation support from senior

decision makers helps to: (1) strengthen the evaluation culture; (2) by making certain evaluation activities a priority, senior managers encourage their teams to get more involved in evaluation; and (3) when senior managers support evaluation, they tend to be more receptive to evaluation recommendations. The main expectations about intended causal processes underlying this message are depicted below.

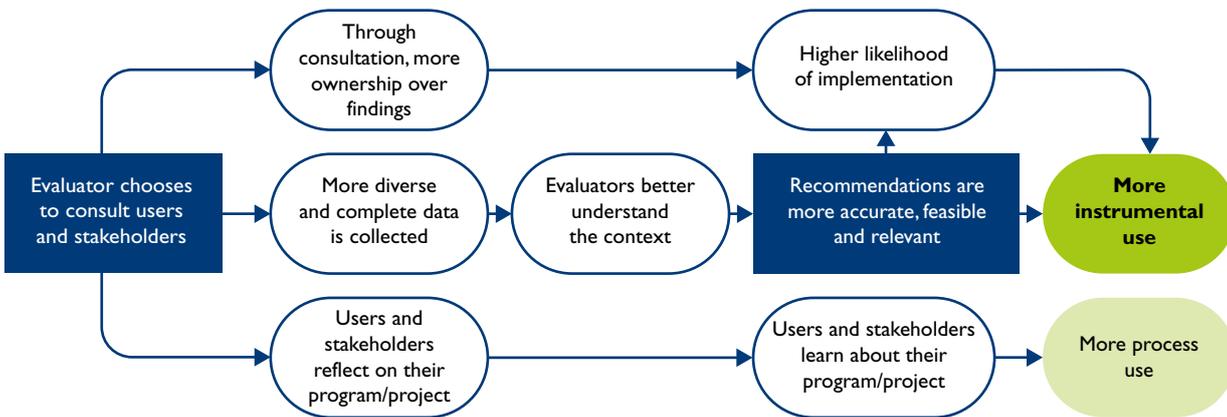
Message 3: Evaluators need to ensure recommendations are feasible and relevant

To increase instrumental use, evaluations should generate relevant and feasible recommendations. Indeed, the higher the quality of recommendations, the more likely that they

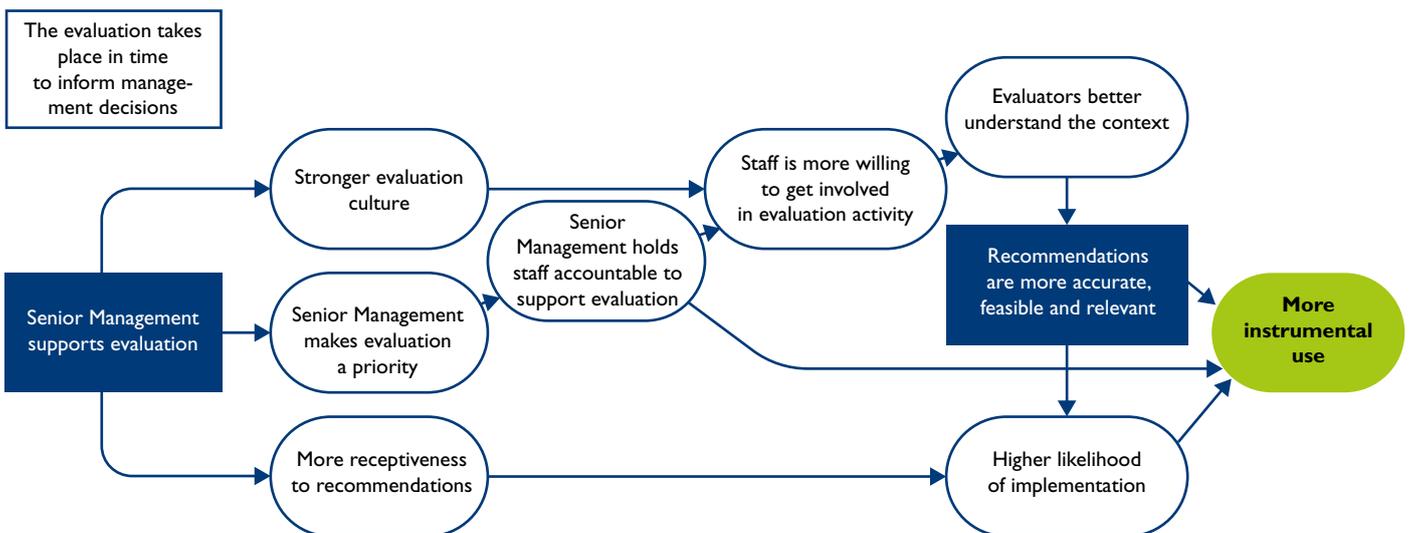
will be accepted and implemented, increasing the potential use of an evaluation. However, proposing recommendations of high quality is not easy and depends on several factors. Most importantly, the presence of: (1) user/stakeholder involvement; (2) a high-quality evaluation methodology; and (3) high-quality evaluators. The main expectations about intended causal processes underlying this message are depicted below.

Message 4: Independent evaluations must make efforts to capture organizational realities

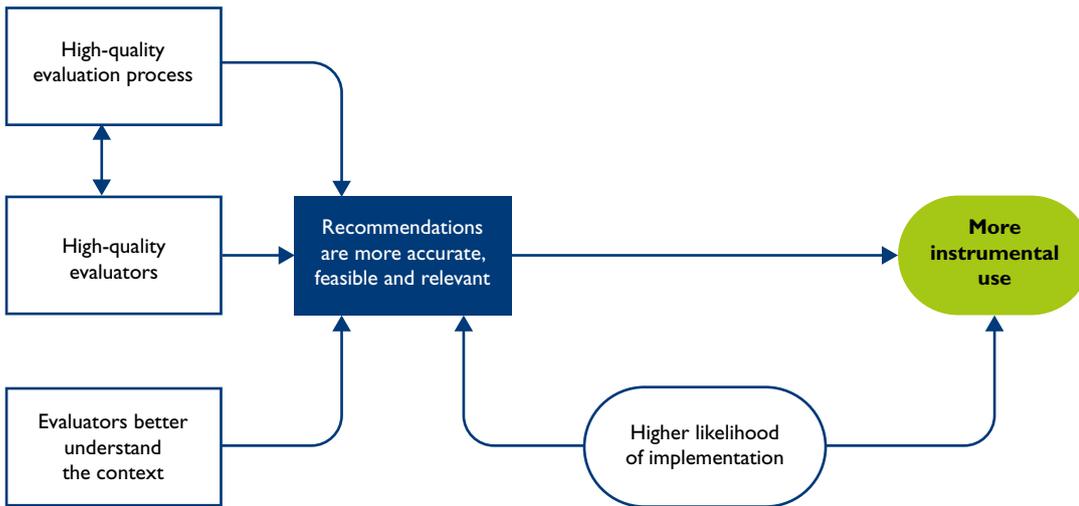
Most professional evaluators would agree that independence is a useful and fundamental principle of evaluation. Across the world,



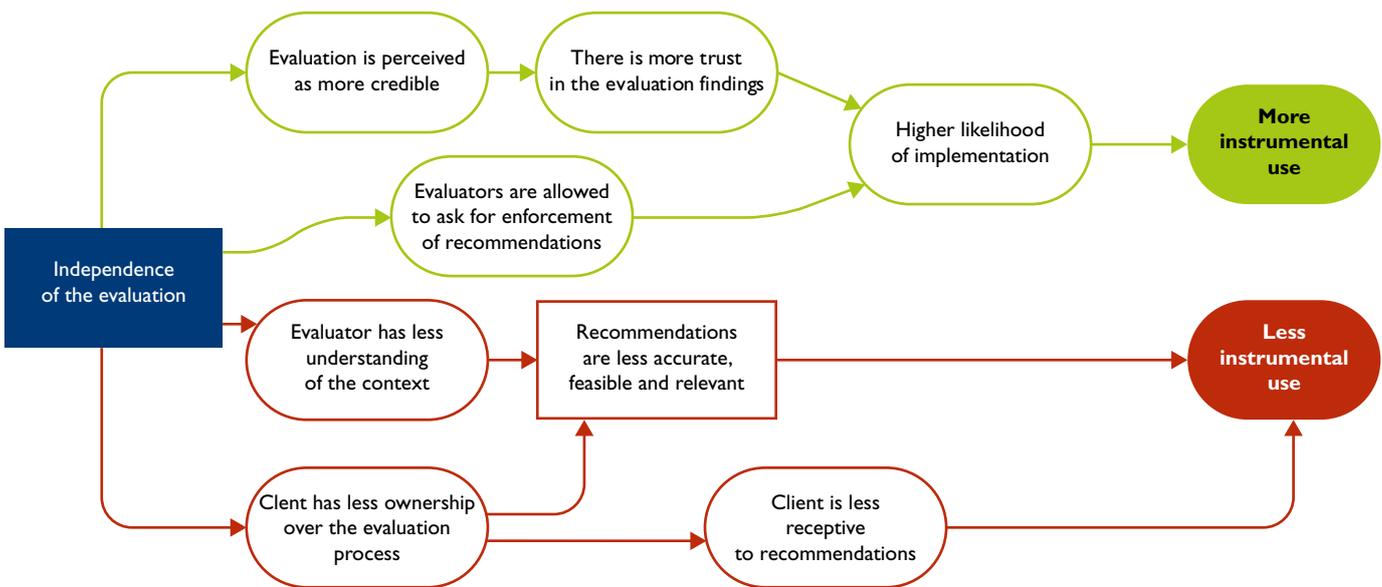
Message 1: Users and stakeholders should be involved in and consulted throughout the evaluation process.



Message 2: The support of senior decision makers is key, and so is their commitment to the implementation of the recommendations.



Message 3: Evaluators need to ensure recommendations are feasible and relevant.



Message 4: Independent evaluations must make efforts to capture organizational realities.

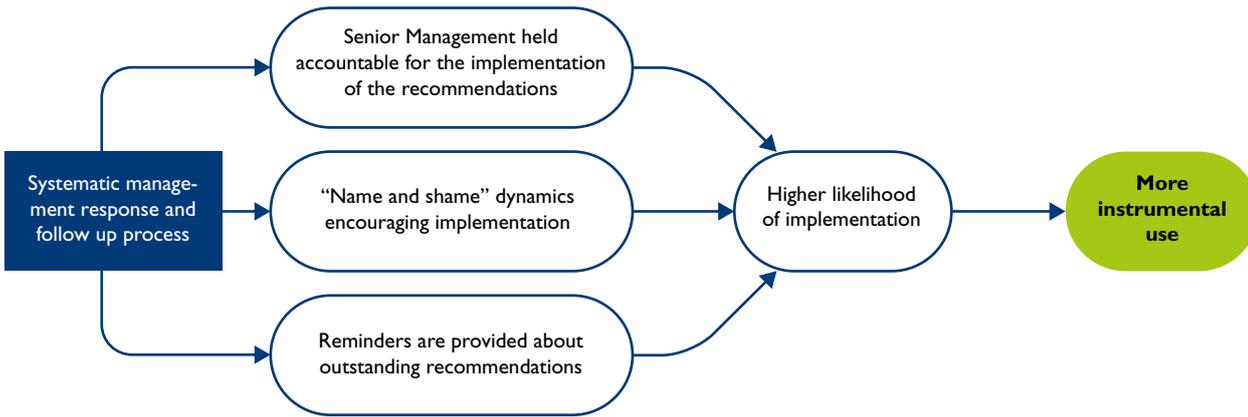
Independence appears to be gaining momentum as a key criterion for evaluation quality. It is a core principle of the OECD-DAC Quality Standards for Development Evaluation, and international organizations – including the UN – are increasingly setting up independent evaluation offices. Independence enhances the impartiality and credibility of evaluations, thereby increasing the stakeholder confidence in, and support for, the recommendations coming out of evaluations.

However, agencies moving towards greater independence face the risk of increasing the distance between the evaluation function and the implementing organization, which may reduce the evaluators' understanding of the context they are evaluating, thereby making recommendations less feasible and relevant. To address this challenge extra effort needs to be undertaken to help evaluators understand the organizational constraints and challenges that might influence the feasibility

of recommendations. The main expectations about intended causal processes underlying this message are depicted below.

Message 5: Management responses and follow-up processes must take place and be adequately supported

In the United Nations system, the most important mechanism for the implementa-



Message 5: Management responses and follow-up processes must take place and be adequately supported.

tion of evaluation recommendations is the management response, as well as the associated follow-up process. Many United Nations organizations have standardized, systematic ways to produce management responses. Management responses are clearly contributing to use as they form the basis for systematic follow-up processes of recommendation implementation. The level of implementation of recommendations can then be published and shared across an organization, which may act as an incentive for future evaluation use. The main expectations about intended causal processes underlying this message are depicted below.

Message 6: Sharing findings would enable cross-organizational learning and use

Evaluation use can be enhanced if evaluations and their findings are actively disseminated.

Evaluation practitioners should be encouraged to make a conscious effort to develop findings and lessons that are relevant and applicable beyond the narrow context of the evaluand and its stakeholders. The main expectations about intended causal processes underlying this message are depicted below.

Conclusion

Evaluation use in function of improved accountability and learning constitutes an ongoing challenge in the UN system. In addressing this challenge, the mechanisms presented above can help evaluators and other stakeholders to better understand the mechanisms underlying evaluation use that may be at work. Some of these mechanisms, which are largely in line with the broader literature on evaluation use, are fairly straightforward and can be relatively easily put in place. In other cases, reality is much more complex.

While one could question particular assumptions about intended use, the bigger question is about how to generate the conditions for these processes to actually take place. Learning from successes and failures, we need to look deeper into the contextual factors that enable and constrain the above mentioned processes to actually materialize.

References

Herbert, J.L. (2014). Researching evaluation influence: A review of the literature. *Evaluation Review*, 38(5), 388–419.

UNEG (2014). *UNEG Strategy 2014–2019*. Retrieved 8 January 2016, from <http://www.unevaluation.org/document/detail/1459>.



Message 6: Sharing findings would enable cross-organizational learning and use.

EVALUATING IFAD’S SUPPORT TO RURAL DEVELOPMENT IN FRAGILE STATES AND SITUATIONS

Oscar A. Garcia, Ashwani K. Muthoo and Fabrizio Felloni

More than 1.2 billion people live in fragile states and situations (i.e. within countries not formally classified as fragile states). People living in fragile states and situations generally exhibit lower economic growth rates, poorer socio-economic indicators and weaker human development indicators. Approximately 40 per cent of the extremely poor (500 million people) live in countries or situations classified as fragile.

As of 2014, 48 Member States of IFAD were classified as fragile: they accounted for about half of the total recipient countries included in the Ninth Replenishment of IFAD's Resources (2013–2015). In 2014 the Independent Office of Evaluation of IFAD conducted a corporate level evaluation on IFAD's engagement in fragile and conflict-affected states and situations (IFAD, 2015).

The evaluation explored IFAD's work over a 10-year period from 2004 until the end of 2013.

IFAD's operations focus on rural poverty reduction, rural development and transformation rather than fragility *per se*. However, a situation of fragility shapes the opportunities and constraints that IFAD faces at the operational and strategic level. Conversely, contributing to rural development may improve a fragile situation, particularly if localized. In a recent contribution, Allouche & Lind (2013) make the case for better integration of development and security interventions. The authors point out the special challenges arising from fragmentation of power and authority and argue that specialized knowledge and skills are needed to understand local realities and power dynamics.

One of the constraints faced by this evaluation was the absence of a common international classification of fragile countries and situations. Moreover, IFAD did not have a unified policy document dealing with fragility. The definition of fragility was formulated in 2006, applied to fragile states (but not fragile sub-national situations within states classified as non-fragile) and had not been revisited since then. As a result, the definition did not adequately reflect recent policy debates (e.g. key principles stemming from the Busan Partnership Agreement).

Another challenge was to conceptualize and capture the linkages between IFAD's work and fragile states and situations. The evaluation prepared a schematic results chain (Figure 1). For each step in the results chain, key assumptions were identified. Although it did

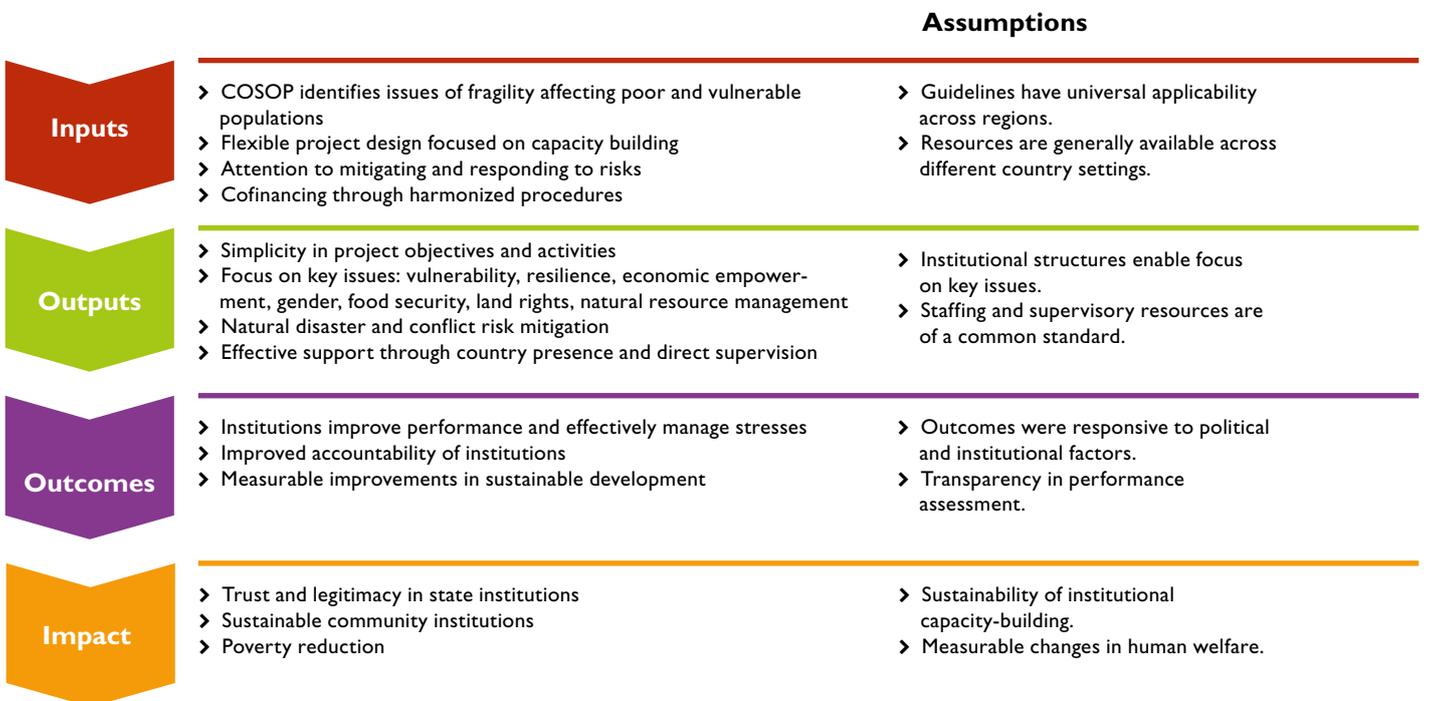


Figure 1. Notional results chain for IFAD's engagement with fragile states*

*The results chain draws on the approach taken by the Independent Evaluation Group at the World Bank in the 2013 Evaluation of IDA's support to Fragile States (World Bank, 2013). Source: IFAD (2015).

not correspond to a full-fledged “theory of change”, it made explicit the tacit approach to fragility delineated in several IFAD documents and constituted a useful structure for the evaluative enquiry. It highlighted the importance of IFAD’s strategy and project design responding to the country fragility context; the nature of project design and implementation support; and outcomes that reflect the needs of a fragile state or situation. Each of these was examined during the evaluation.

A mixed methods approach was adopted and encompassed the following aspects: (i) review of definitions, concepts and approaches to fragility and conflict; (ii) review of IFAD’s relevant policy framework; (iii) analysis of portfolio performance documentation, including 42 country strategies (and including independent evaluations); (iv) a web-based questionnaire sent to all IFAD country program managers and staff in country offices; and (v) ten country visits-based case studies to develop perspectives from the field in all five IFAD geographic regions.

The ten countries were selected so as to represent cases of: (i) persistent fragility; (ii) volatile fragility (moving out and back into fragility); (iii) graduation from fragility; (iv) ‘non-fragility’ at the overall country level but with fragile situations at the sub-national level. Among other things, the evaluation assessed the drivers of fragility in the ten case studies.

The evaluation found that the performance of IFAD-funded projects in countries that had always been classified as fragile was lower than in countries that have moved in and out of fragility or were never classified as fragile. Achievements in promoting gender equality and women’s empowerment had not improved in countries that had always been fragile, whereas they showed improvement in countries that were only partially or never fragile. There were significant improvements in countries that had always been fragile in

terms of overall project achievement, project effectiveness, IFAD’s performance as a partner, and rural poverty impact.

The evaluation identified a number of drivers contributing to better performance of projects in fragile states. These included: the transformation of IFAD’s operating model which included a move to direct supervision of projects since 2007 and more thorough portfolio monitoring and management to improve the implementation of projects at risk; the opening of IFAD country offices, with close to half of IFAD country offices being located in fragile states.

On the other hand, the evaluation pinpointed some of the key explanatory factors for weaker performance. First of all, the context analyses in strategic and project documents did not sufficiently examine drivers of fragility. This resulted in over-ambitious objectives and complex project designs, adding an additional stress factor to an already constrained institutional capacity. Second, the evaluation found that there was limited availability of skills in (and provision of training for) staff working in fragile situations. Also, budgets for analytical work, design, supervision, implementation support, and self-evaluation were not adequate for fragility situations.

The 2011 World Bank Development Report on Conflict, Security, and Development (World Bank, 2011) highlighted the need for development efforts to go beyond institutional fragility and socio-political instability to target the root causes and drivers of conflict and fragility and break cycles of violence. In line with this perspective, the evaluation found that IFAD has a critical role to play in fragile and conflict-affected states and situations in promoting sustainable inclusive development and rural transformation.

A very large number of people live in severe poverty in such contexts. As the only multi-lateral development organization that focuses exclusively on smallholder agriculture de-

velopment in rural areas, IFAD has a unique responsibility to support local production and livelihoods systems in fragile situations, and help poor rural people improve their incomes, nutrition levels, food security and well-being.

According to the evaluation, good work has been done in recent years to promote improved approaches to gender equality and institution building. Although they have not yet had a notable impact on IFAD’s wider portfolio, they provide a benchmark for ways of working that need to be taken up in other areas of work such as the reintegration of former combatants in the rural economy, access to land and land tenure, and youth employment. In order to achieve this IFAD will, however, need to further adapt and refine its approaches.

References

- Allouche, J. & Lind, J. (2013). *A New Deal? Development and Security in a Changing World*. Institute of Development Studies (IDS). Brighton: IDS.
- IFAD (2015). *Corporate Level Evaluation of IFAD's Engagement in Fragile and Conflict-affected States and Situations*. Retrieved 11 January 2016, from http://www.ifad.org/evaluation/public_html/eksyst/doc/corporate/fragile/cle_fragile.pdf
- World Bank (2011). *World Development Report 2011: Conflict, Security, and Development*. Washington, DC: World Bank.
- World Bank (2013). *World Bank Group Assistance to Low- Income Fragile and Conflict-Affected States. An Independent Evaluation*. Independent Evaluation Group. Washington, DC: World Bank.

EVALUATING STANDARD-SETTING WORK IN UNESCO: CULTURE CONVENTIONS

Barbara Torggler

Setting standards in UNESCO's areas of competence has been part of the organization's work since the 1950s. Several of UNESCO's culture conventions are among the most visible and best known standard-setting instruments¹ of the organization. A relatively recent survey undertaken by UNEG revealed that within the UN system normative (standard-setting) work is usually assessed as a component of wider evaluations, but seldom as a stand-alone exercise (UNEG, 2012). Not many evaluations develop a comprehensive perspective on the normative work undertaken in the framework of a particular legal instrument: the support provided to the development of the instrument, its ratification; its integration at the level of policy and legislation of Parties; its implementation at the program level through the provision of manuals and guidelines, capacity building, technical support and advocacy; and, related intergovernmental processes at different levels.

Evaluating the standard-setting work in culture

In 2012, UNESCO's Internal Oversight Service (IOS) initiated a series of evaluations covering four culture conventions (UNESCO

2013, 2014a, 2014b, 2014c). Conducting comprehensive evaluative exercises for several instruments was expected to be challenging given their broad scope and the inherent complexity of the evaluands.² While most of the conventions have one main objective, this objective is usually very broad, containing numerous, not always well-defined sub-objectives. Implementation procedures are well-defined at the global level, but not necessarily at national levels. Conventions have a clear start date, but no clear end date, and their orientation is long-term. And most importantly, the substantive issues underlying the conventions are highly complex as well, since culture is closely intertwined with questions of identity, social cohesion, gender, and sustainable development.

Each of the conventions has its own governing bodies; funding, implementation and monitoring mechanisms; and capacity building programs. Governing bodies are assisted by the UNESCO Secretariat. In addition, conventions involve many other stakeholders at different (international, national, local) levels, including the Executive Board of UNESCO, government and civil society organizations, private sector actors, academia, expert communities, donors, etc. Implemen-

tation and funding agencies involved are also diverse.

Causality and change processes are not clearly articulated for any of the conventions.³ Often these are non-linear and interconnected. As a consequence, the level of agreement and clarity on appropriate actions to be taken can be low. Political considerations often play a role. And last but not least, contextual factors have a major influence on the implementation of the conventions, as well as their potential effects. These include activities as well as phenomena (mining, infrastructure development, tourism, illegal trafficking of cultural objects, climate change, etc.) that threaten the safeguarding of cultural heritage. In addition, there are often competing priorities that divert attention and resources from culture to other sectors.

In light of the complexity of the exercise, the limited availability of financial resources, and an overall evaluation time frame of only about 1.5 years, it was decided that the evaluation would focus on only four of the six culture conventions. For each convention, it looked at ratification and implementation at the level of policy and legislation of Parties to the convention. In addition, on the basis of the same

		1954 Convention	1970 Convention	1972 Convention	2001 Convention	2003 Convention	2005 Convention
Audit	Working methods						
Evaluation	Ratification						
	Policy / legislation level implementation						
	Program level implementation						
	UNESCO support mechanisms						

Table 1. Scope and coverage of the evaluation of UNESCO's standard-setting work in culture.

considerations as well as indications of stakeholder demand, the evaluation focused on two conventions (1970, 2003) in more detail, also covering the implementation at program level and the relevance and effectiveness of UNESCO's support mechanisms. The evaluation (consisting of four independent evaluation exercises) was accompanied by an audit of the working methods of all six culture conventions. The exercise concluded with an analysis of cross-cutting issues (see Table I).

Given the complexity of causality between the different levels of standard-setting work and the absence of any clear articulation of change processes for the conventions, theories of change were reconstructed for the 1970 and 2003 conventions. For these two conventions, the evaluation methodology was designed around a multi-level framework comprising three levels: ratification; integration of the provisions of the conventions into national/regional legislation, policies and strategies (policy development level); implementation of the legislation, policies and strategies at national level (policy implementation level). A multi-level purposive sampling strategy was developed which started out from a sample from the broad population of all countries at the ratification level, and gradually narrowed down to sampling from smaller populations of countries with certain levels of policy development and implementation related to a particular convention (or a lack thereof). The multi-level theories of change for the two conventions and corresponding data collection frameworks constitute a good example of "unpacking" an intervention into evaluable parts as described in Bamberger et al. (2016).

The evaluations drew upon a large number of different knowledge sources: literature, samples of policy/legislation of Parties, implementation reports, interviews with different types of stakeholders, survey results, and field observation. Convention secretariats were involved throughout the evalu-

ation process and stakeholder groups were consulted both at the global level (including members of the governing mechanisms of the conventions) and at national levels. Evaluation recommendations responded to the needs and challenges of relevant stakeholders identified at each of these levels (secretariats, governing mechanisms, Parties).

Final remarks on evaluation use

Three of the evaluations and the audit came *at the right time*, but not necessarily for the same reasons. Overall, the evaluations were used to inform decision-making processes, which were quite influenced by the organization's budget cuts and the resulting need for more efficiency and focus. Furthermore, 2003 Convention stakeholders were ready to take stock after almost 10 years of implementation, and to review and reorient their ways of working. For the 1970 Convention, a new governing mechanism had just been established, which created an opening for change. For one Convention (1972) the timing was *not right* for the convention secretariat; its capacities were stretched and it did not see the added value of the evaluation given that several (performance) audits had been conducted over the past years. In hindsight, IOS should have further explored the usefulness of including the 1972 Convention in the evaluation.

Overall, evaluation use (instrumental, process, conceptual and even political use) seems to be higher where extensive consultation and involvement of stakeholders had taken place. The fact that stakeholders do not ask for an evaluation does not mean they will not find it useful later on. This essentially depends on how the evaluation is conducted and how stakeholder consultation is organized. Making recommendations to Parties to the conventions works (even for an internal evaluation function such as IOS) as long as governing mechanisms assume ownership and follow up on implementation. ■

References

- Bamberger, M., Vaessen, J. & Raimondo, E. (2016). *Dealing with Complexity in Development Evaluation: A Practical Approach*. Thousand Oaks: SAGE.
- UNEG (2012). *UNEG Task Force on Evaluation of Normative Work – Survey Analysis*. New York: UNEG.
- UNESCO (2013). *Evaluation of UNESCO's Standard-setting Work of the Culture Sector: Part I – 2003 Convention for the Safeguarding of the Intangible Cultural Heritage*. Retrieved 5 February, from <http://unesdoc.unesco.org/images/0022/002230/223095E.pdf>.
- UNESCO (2014a). *Evaluation of UNESCO's Standard-setting Work of the Culture Sector: Part II – 1970 Convention on the Means of Prohibiting and Preventing the Illicit Import, Export and Transfer of Ownership of Cultural Property*. Retrieved 5 February, from <http://unesdoc.unesco.org/images/0022/002269/226931E.pdf>.
- UNESCO (2014b). *Evaluation of UNESCO's Standard-setting Work of the Culture Sector: Part III – 1972 Convention Concerning the Protection of the World Cultural and Natural Heritage*. Retrieved 5 February, from <http://unesdoc.unesco.org/images/0022/002269/226922E.pdf>.
- UNESCO (2014c). *Evaluation of UNESCO's Standard-setting Work of the Culture Sector: Part IV – 2005 Convention on the Protection and Promotion of the Diversity of Cultural Expressions*. Retrieved 5 February, from <http://unesdoc.unesco.org/images/0022/002269/226932E.pdf>.
- UNESCO (2015). *Reshaping Cultural Policies – 2005 Convention Global Report*, Retrieved 5 February, from <http://unesdoc.unesco.org/images/0024/002428/242866E.pdf>.

1 The evaluation worked with UNEG's definition of normative (standard-setting) work (2013).

2 Reference is made to the dimensions of complexity, as discussed in Bamberger et al. (2016).

3 This has changed since. A monitoring framework was recently established for the 2005 Convention (UNESCO, 2015). Work is also planned to develop a results framework for the 2003 Convention.

EVALUATION AND THE GLOBAL ENVIRONMENT: CHALLENGES AND OPPORTUNITIES

Juha I. Uitto

Despite increased attention to (and growing investments into) issues such as climate change mitigation, biodiversity conservation and degradation of ecosystems, most global environmental trends continue to deteriorate. Much of this stems from the nature of the environment as a global public good. The benefits accrue to everyone but destructive activities seldom result in costs to those causing them. The Global Environment Facility (GEF) serves as financial mechanism for several international environmental conventions, including those on biodiversity, climate change, desertification, persistent organic pollutants, and mercury.

As the largest single public source of funding for these issues, since 1991 the GEF has provided US\$14.5 billion in grants and mobilized US\$75.4 billion for almost 4,000 projects around the world. Still, public funding to global environmental projects and programs is dwarfed by financial flows to environmentally destructive subsidies, whether to fossil fuels, unsustainable agriculture or other environmentally hazardous practices. Environmental policies, strategies, programs and projects must be focused and effective to make a difference. It is not enough just to evaluate that the projects and programs achieve their outputs; ensuring that we are actually helping the global environment and the people depending on it is essential to ascertain that we are contributing to sustainable development.

This article addresses global environment evaluation issues in light of experience gained by the GEF Independent Evaluation Office. Evaluation in the GEF is faced with specific challenges, but also opportunities to explore multiple approaches and methods to gauge the success of the work and for learning what works, under what circumstances. GEF actions focus primarily on three areas: providing knowledge and information to countries and partners to address global environmental challenges; supporting institutional capacity, including policy, legal and regulatory frameworks, and governance

structures and arrangements; and supporting implementation of programs and projects through technologies, financial mechanisms, and other means.

These policy interventions are intended to lead to behavioural changes amongst the various actors and to promote broader adoption of environmentally sound practices through mainstreaming, replication, scaling-up and market change. In practical terms, evaluation must be able to assess the behavioural changes and broader adoption as a result of GEF interventions, and to get to the actual impact on the environmental status and stress reduction.

This task is complicated by challenges specific to environmental evaluation, such as the different time horizons and geographical scales for changes in natural and human systems, compelling us to adopt a two-system evaluation (Rowe, 2012). Project timeframes are often too short for environmental impacts to appear. There are also differing spatial scales, as environmental phenomena do not adhere to political and administrative boundaries that are inherent in human systems and most projects. Watersheds are often shared by different administrative and jurisdictional units, including countries, which water, air and migratory species cross.

Finally various stakeholder groups often hold divergent values. In these circumstances, theory-based evaluation is a useful tool to understand the diverse spatial, temporal, organizational and value scales. However, modelling theories of change using the successionist causality concept (i.e., the traditional results chain) is not satisfactory as the programs and projects operate in complex environments where the intervention is part of a larger system.

For evaluations, it is important to understand the various elements in such systems, their boundaries, components, interactions and emergent properties (Garcia & Zazueta, 2015). Such an understanding is necessary for

us to define the evaluation scope, appropriate units of analysis, and suitable approaches and methodologies in an iterative manner. Before starting a major evaluation, we also build upon existing scientific knowledge on especially the behaviour of the natural systems.

An example of when defining the system boundaries was essential concerns the Impact Evaluation of the GEF in the South China Sea and Adjacent Areas (GEF, 2012), which was undertaken to assess the environmental and institutional results of two decades of GEF investments through multiple projects in the area. The evaluation found that, while GEF was only one actor amongst many in the region, it had played a unique role in enabling contributions to larger regional processes, linking initiatives at multiple scales.

To get to the actual on-the-ground impacts, the recently completed Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems (GEF, 2015) sought to answer three broad questions: (i) what have been the impacts and contributions of GEF support (positive or negative, intended or unintended) on biodiversity conservation in protected areas and adjacent landscapes; (ii) what have been the contributions of GEF support to the broader adoption of biodiversity management measures at the country level and what were the key factors at play; and (iii) which GEF supported approaches and contextual conditions are most significant? Adopting a theory of change approach, the evaluation assumed that the presence of good governance, effective protected area management, and community engagement together lead to improvements in biodiversity conservation.

The evaluation utilized a mix of quantitative, qualitative and spatial methods in data collection and analysis at three specific levels: (i) portfolio analysis (including databases and project-level terminal evaluations); (ii) global analysis (using methods such as principal component analysis, propensity score

matching, linear regression, and remote sensing and GIS analysis); and (iii) case studies in seven countries across three regions covering 17 GEF-supported protected areas and 11 non-GEF protected areas (using qualitative data analysis and synthesis tools, as well as remote sensing).

Despite significant information gaps on GEF support to protected areas, limited global time series data, and difficulties in estimating counterfactuals, the evaluation was able to identify directions and patterns regarding GEF's contributions towards biodiversity conservation and interactions with the larger social-ecological system. The evaluation was able to conclude that GEF support has contributed to biodiversity conservation, especially through reducing habitat loss, as well as building capacities and contributing to large-scale change in biodiversity governance in countries.

As a key element in the current 6th replenishment period, the GEF is embarking on integrated approach pilots to achieve broader

impact. The strategy is intended to address the drivers of environmental degradation and to support innovative and scalable activities. The three programs – focused on Fostering Sustainability and Resilience for Food Security in Sub-Saharan Africa, Sustainable Cities, and Taking Deforestation out of Commodity Supply Chains – all promote integrated approaches to address global environmental problems. One of the challenges the GEF must confront is the tension between global, national and local priorities. While programming will seek win-win situations, there will inevitably be trade-offs between economic and environmental priorities. For global environmental interventions to be successful they must contend with the livelihoods and vested interests in productive and extractive activities.

The implications for evaluation are equally clear. We must be able to evaluate effectively where economic and social development and environmental protection meet. For this, we must further hone our approaches and methodologies to address issues pertaining to the

combined human and natural systems. This touches upon the heart of the 2030 Agenda for Sustainable Development and will be crucial for evaluating the new Sustainable Development Goals.

References

Garcia, J.R. & Zazueta, A. (2015). Going beyond mixed methods to mixed approaches: A systems perspective for asking the right questions. *IDS Bulletin* 46(1), 30–43.

GEF (2012). *The GEF in the South China Sea and Adjacent Areas*. Washington, DC: Global Environment Facility Evaluation Office.

GEF (2015). *Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems*. Washington, DC: Global Environment Facility Independent Evaluation Office.

Rowe, A. (2012). Evaluation of natural resource interventions. *American Journal of Evaluation* 33(3), 384–394. ■

THE AUTHORS

George Bartsiotas



has been working for more than 30 years in management and senior executive positions in the public sector and in international organizations. Before becoming an Inspector at the UN Joint Inspection Unit (JIU), he held positions in policy-making, management and administration, and in a range of areas of oversight including audits, evaluations, inspections, and management reviews. He served as senior consultant and adviser to UN system organizations; was Director of Internal Oversight at the OSCE; served as Counsellor and Director of Resource Management at the US Mission to Austria; was Deputy High Representative at the Office of the High Representative (OHR) in Bosnia and Herzegovina; and was posted at US Embassies as a member of the senior foreign service. He is a CGAP, a CFE, a Quality Assurance Examiner, and a Registered Professional Engineer.

Fabrizio Felloni



is a Lead Evaluation Officer at the Independent Office of Evaluation (IOE) of the International Fund for Agricultural Development (IFAD), since July 2010. Previously, he was Evaluation Specialist at the Evaluation Office of the United Nations Development Programme (New York) for three years and Evaluation Officer at IFAD for six years. He has worked on international development evaluation since 2000, leading and participating in over a dozen country programme and project-level evaluations in Africa, Asia, Latin America and Eastern Europe. He holds a Master's degree in Agricultural Economics from Washington State University (US) and a Master-equivalent degree in Social and Economic Sciences from Bocconi University (Italy). He has been the author and co-author of several articles published in peer-reviewed journals.

Oscar A. Garcia



is the Director of the Independent Office of Evaluation (IOE) of the International Fund for Agricultural Development (IFAD). Prior to joining IFAD, he served as Head of the advisory services at the United Nations Environment Programme, Division of Technology, Industry and Economics (Paris). Previously, he was senior evaluation advisor at the Evaluation Office of UNDP (New York) where he led significant programmatic and thematic evaluations. He co-chaired the UNEG task forces on Country Level Evaluations, Capacity Development in Evaluation, and Harmonization of Evaluations in the UN system. Mr Garcia holds two Masters' degrees in Organizational Change Management from the New School University, NY (US), and an MBA from the Catholic University of Bolivia in association with the Harvard Institute for International Development. He has authored several evaluation reports and papers on development evaluation.

Jos Vaessen



(PhD Maastricht University) is Principal Evaluation Specialist at the Internal Oversight Service of UNESCO and lecturer at Maastricht University. After completing his MSc in 1997 (Wageningen University) and prior to starting his current position at UNESCO in 2011, he has been involved in research, teaching and evaluation activities in the field of international development at Antwerp University and, more recently, Maastricht University. Over the last fifteen years or so, he has worked for several multilateral and bilateral international organizations mostly on evaluation-related assignments. He has been author of several internationally peer-reviewed publications, including three books. His most recent book is: *Dealing with complexity in development evaluation: a practical approach* (2016, with Michael Bamberger and Estelle Raimondo).

Ashwani K. Muthoo



currently holds the position of Deputy Director in the Independent Office of Evaluation (IOE) of the International Fund for Agricultural Development (IFAD). He has extensive experience in international cooperation, agriculture and rural development, management and development evaluation. During more than 25 years, Mr Muthoo conducted significant evaluations that have contributed to IFAD's institutional transformation and improved development effectiveness. His work includes corporate-level evaluations on IFAD's institutional and operational efficiency; agriculture in Africa jointly undertaken with the African Development Bank; the field presence pilot programme; gender; private sector engagement; fragile states; IFAD replenishments; and innovation and scaling-up. He is also responsible for preparing the Annual Report on Results and Impact of IFAD Operations (ARRI), IOE's annual flagship report.

Sukai Prom-Jackson



is an Inspector with the Joint Inspection Unit of the UN System (JIU) (2013–2017). Her work at the JIU is focused on strengthening the development of systems to enhance the provision of strong evidence to support system-wide policy making in support of both organizational and development effectiveness of the UN system. She has over 25 years of senior level professional experience and leadership in the management and conduct of research and evaluation, in policy formulation and strategic planning, in programme development, and in the facilitation of learning. She has worked for the UNDP, World Bank, US government, universities and research centers in the US, and The Gambia government. Her direct work experience has covered over 80 countries.

Michael Bamberger



worked on urban community development in Latin America for over a decade and in the evaluation, gender and training departments of the World Bank. He has consulted and taught on development evaluation with UN, multilateral and bilateral agencies and international NGOs. He advises the Evaluation Office of the Rockefeller Foundation and is a member of the UNDP Independent Evaluation Office International Advisory Panel. Recent publications include: *RealWorld Evaluation: Working under budget, time, data and political constraints* (2012, with Jim Rugh and Linda Mabry), *How to design and manage equity-focused evaluations* (2011, with Marco Segone) and *Dealing with complexity in development evaluation* (2016, with Jos Vaessen and Estelle Raimondo).

Marco Segone



is Director, Independent Evaluation Office, at UN Women; Chair, United Nations Evaluation Group (UNEG), the network of Evaluation Offices of 46 UN agencies; and co-founder of EvalPartners, the global partnership for national evaluation capacities. Previously, he was responsible for the decentralized evaluation function as well as the national evaluation capacity development portfolios at the UNICEF Evaluation Office; Regional Chief, Monitoring and Evaluation in the UNICEF Regional Office for Europe and Central Asia; Regional Office for Latin America and the Caribbean; Brazil Country Office, and Niger Country Office. Marco also worked in international NGOs in Albania, Pakistan, Bangladesh, Thailand and Uganda. He has authored numerous publications including *Evaluation for equitable development results* and *How to design and manage equity-focused evaluations*.

Robert Stryk



is the vice-chair for use of evaluation in the United Nations Evaluation Group, leading research and outreach advocating better use of evaluations. He is also the Chief of

the Evaluation Division in the United Nations Relief and Works Agency for Palestine Refugees. During his 20 years career in development he has focused on evaluation and strategic planning.

Barbara Torggler



works as a Principal Evaluation Specialist at UNESCO's Internal Oversight Service, where she is in charge of leading and managing strategic evaluations of UNESCO's policies and programmes. Recent work includes several evaluations of the standard-setting work of UNESCO's culture and education sectors (conventions), and an evaluation of UNESCO's work on Culture and Sustainable Development. Barbara has a background in social and cultural anthropology, economics and languages, and 20 years of work experience on issues related to international cooperation and sustainable development, involving programme design, management and evaluation. She has had assignments all over the world, including for UNESCO, UNODC, UNDP, a bilateral donor agency and an international NGO. She has been a university lecturer, facilitator and coach.

Juha Uitto



is Director of the Global Environment Facility (GEF) Independent Evaluation Office. Since 1999, he has worked on evaluation with the GEF and UNDP, including as Evaluation Adviser and Deputy Director at the UNDP Independent Evaluation Office. He has conducted and managed a large number of programmatic and thematic evaluations at the country, regional and global levels, especially in Asia and Africa. His career has combined positions in international development and academia and he has held positions at the UN University, where he coordinated the environment and sustainable development research and training programme, the Nordic Africa Institute and IFAD. Educated at the universities of Helsinki and Lund, Juha earned his PhD in Social and Economic Geography.

GUIDANCE TO CONTRIBUTORS

Brevity and nimbleness are the hallmarks of Connections: published articles are normally 800–1,200 words long. Even shorter contributions (news items; opinion pieces; book reviews and letters to the editor) are accepted with a view to promote debate and connect evaluators within Europe and beyond. While Connections is not a peer reviewed publication only articles that add to knowledge about the theory, methods and practices of evaluation should be submitted.

Contributions that highlight European values and evaluation practices are given

priority but Connections also reaches out beyond Europe to the international evaluation community and favours articles reflecting a diversity of perspectives. Within the limits of copyrights agreements articles that summarize in a cogent way the substantive content of published (or to be published) studies, papers, book chapters, etc. are welcome (with suitable attribution).

Individuals or organizations wishing to sponsor special issues about an evaluation theme or topic of contemporary interest should contact the EES Secre-

tariat (secretariat@europeanevaluation.org). Such special issues usually consist of 6–8 articles in addition to a guest editorial. A Presidential letter may be included. The guest editor(s) are responsible for the quality of the material and the timeliness of submissions. The regular editorial team ensures that special issues meet 'Connections' standards and takes care of copy editing.

To facilitate copy editing, authors are encouraged to use end notes rather than footnotes and to use the APA style guide for references. Here are some examples:

For books: Bergmann, I. (1997). Attention deficit disorder. In *The new Encyclopedia Britannica* (Vol. 26, pp. 501–508). Chicago, IL: Encyclopedia Britannica.

For journal articles: Rindermann, H., & Ceci, S. J. (2009). Educational policy and country outcomes in international cognitive competence studies. *Perspectives on Psychological Science*, 4(6), 551–568. doi:10.1111/j1745-6924.2009.01165.x

Website: United States Environmental Protection Agency. (2007, May 4). Climate Change. Retrieved 12 June 2014 from the Environmental Protection Agency website: <http://www.epa.gov/climatechange>.

In-text reference: (United States Environmental, 2007).

www.ees2016.eu

12th EES Biennial Conference

Evaluation Futures in Europe and beyond Connectivity, Innovation and Use

MECC Maastricht, the Netherlands, 26–30 September, 2016



EUROPEAN
EVALUATION
SOCIETY

www.ees2016.eu

