UNEG
United Nations Evaluation Group

United Nations contributions to national evaluation capacity development and the evolution of national evaluation systems

An overview of implementation of General Assembly Resolution 69/237

2022

Cover photo under license from Shutterstock.com

UNEG
United Nations Evaluation Group

United Nations contributions to national evaluation capacity development and the evolution of national evaluation systems

An overview of implementation of General Assembly Resolution 69/237

2022

Acknowledgements

This report has been prepared by the United Nations Evaluation Group Working Group on National Evaluation Capacity Development, with the support of Professor Ian Goldman, who led the main data collection and analysis underpinning the final report. Natalia Kosheleva drafted a preliminary report on progress in the implementation of United Nations General Assembly Resolution 69/237, which provided the starting point for the present report. Six case studies were prepared, by Ian Goldman (Kenya, Philippines, Sri Lanka), Miguel Angel Lombardo Chico (Costa Rica, Morocco), and Emmanuel David-Gnahoui (Benin). The report could not have been completed without the involvement of a wide range of stakeholders who generously shared their time and ideas with the consultants and the Working Group.

The Working Group is co-chaired by Heather Bryant (UNDP) and Renata Mirulla (FAO). Active members who have contributed to this report include: Carlos Asenjo Ruiz (UNODC), Michala Assankpon (WFP), Martin Barugahare (UN-Habitat), Genny Bonomi (FAO), Hanife Cakici (OIOS), Oyuntsetseg Chuluundorj (UNFPA), Michael Craft (UN Women), Asela Kalugampitiya (UNFPA), Aurelie Larmoyer (FAO), Riccardo Polastro (UNICEF), Carlos Rodriguez Ariza (PAHO), Rica Terbeck-Soine (WFP), and Patricia Vidal Hurtado (ILO).

Contents

Acknowledgements 2

Abbreviations and acronyms 5

Executive summary 7

1. Introduction 15

1.1 Introduction 15

1.2 Structure of the report 16

1.3 Towards a new United Nations General Assembly Resolution 17

1.2.1 Timeline and key milestones 17

1.2.2 What has happened since 2014 19

1.4 Methodology 22

2. National evaluation capacity development 25

2.1 Definitions and concepts 25

2.2 A theory of change for NECD 29

3. United Nations support to NECD and the evolution of national evaluation systems 35

3.1 United Nations support to NECD activities 35

3.2 Overview of the case study countries 39

3.3 Fostering an enabling environment 40

3.3.1 Policy and regulatory environment 40

3.3.2 Role and engagement of stakeholders 43

3.3.3 Evaluation champions and persistent political will 44

3.4 Institutional capacity 46

3.4.1. Evaluation institutions established 46

3.4.2 Institutional mechanisms for evaluation in place 47

3.4.3 Strengthening national evaluation systems through United Nations evaluations 51

3.4.4 Institutional capacity: non-public sector organizations 54

3.5 Individual capacity 56

3.5.1 Developing individual capacity to manage, conduct and use evaluations 56

3.5.2 Strengthening understanding of the role of evaluative evidence 59

3.6 Leading to change: demand, production and use of evaluations for decision-making 61

4. Reflections on the way forward for UNEG 67

4.1 Undertaking United Nations evaluations with a NECD perspective 67

4.1.3 United Nations Sustainable Development Cooperation Framework evaluations and NECD 70

4.2 Direct support to strengthening national evaluation capacity and systems 71

4.2.1 Lessons from the case studies on supporting national evaluation systems 71

4.2.2 Entry points for wider NECD work 73

4.3 United Nations agency engagement in NECD and resources 74

4.4 **Coordination and collaboration** 75

4.4.1 Coordination and collaboration among United Nations agencies 75

4.4.2 Coordination and collaboration beyond United Nations agencies 78

4.4.3 Evaluation and achievement of the SDGs 80

4.5 Adapting evaluation to meet global crises 82

5. Conclusions and recommendations 85

5.1 Conclusions 85

5.2 Recommendations 87

Bibliography 91

Appendix 1. Extracts from relevant United Nations resolutions 95

Appendix 2. Survey of UNEG Members on National Evaluation Capacity Development 98

Appendix 3. Members of Working Group (and showing those interviewed) 104

Appendix 4. People interviewed outside the Working Group 105

Abbreviations and acronyms

3iE	International Initiative for Impact Evaluation
ADB	Asian Development Bank
AfrEA	African Evaluation Association
AME	Moroccan Evaluation Association
BEPPAAG	Bureau of Evaluation of Public Policy and Analysis of Government Action (Benin)
CIMES	County Integrated M&E System (Kenya)
CLEAR	Centre for Learning on Evaluation and Results
CSO	Civil society organization
DBM	Department of Budget Management (Philippines)
DEval	German Institute for Development Evaluation
DFAT	Australian Department of Foreign Affairs and Trade
DFID	United Kingdom Department for International Development
DMEO	Development Monitoring and Evaluation Office (India)
DPME	Department of Planning, Monitoring and Evaluation (South Africa)
DPMM	Department of Project Management and Monitoring (Sri Lanka)
DSWD	Department of Social Welfare and Development (Philippines)
EAPRO	East Asia and Pacific Regional Office
ESK	Evaluation Society of Kenya
FAO	Food and Agriculture Organization of the United Nations
FOCEVAL	Evaluation Capacity Development in Latin America project
GEF	Global Environment Facility
GEI	Global Evaluation Initiative
GIZ	German Technical Cooperation Agency
HEPP	Joint Programme to Strengthen Harmonization and Public Policy Evaluation
IEO	Independent Evaluation Office (UNDP)
IES	Independent Evaluation Section (UNODC)
IFAD	International Fund for Agricultural Development
ILO	International Labour Organization
INCE	National Evaluation Capacity Index
INDH	National Initiative for Human Development (Morocco)
IOM	International Organization for Migration

M&E	Monitoring and evaluation
MED	Monitoring and Evaluation Directorate (Kenya)
MfDR	Managing for Development Results
Mideplan	Ministry of National Planning and Economic Policy (Costa Rica)
MPI	Ministry of Plan Implementation (Sri Lanka)
NDP	National Development Plan (Costa Rica)
NEC	National evaluation capacity
NECD	National evaluation capacity development
NEDA	National Economic and Development Authority (Philippines)
NEP	National Evaluation Policy
NEPF	National Evaluation Policy Framework
NGO	Non-governmental organization
NIMES	National Integrated Monitoring and Evaluation System (Kenya)
OECD	Organization for Economic Cooperation and Development
OIOS	Office of Internal Oversight Services
ONDH	National Observatory of Human Development (Morocco)
PAHO	Pan American Health Organization
PPE	Public policy evaluation
RBM	Result-based management
SDG	Sustainable Development Goal
SIDA	Swedish International Development Cooperation Agency
SLEvA	Sri Lanka Evaluation Association
UMI	Moulay Ismail University (Morocco)
UNDAF	United Nations Development Assistance Framework
UNDP	United Nations Development Programme
UNEDAP	United Nations Evaluation Development Group for Asia and the Pacific
UNEG	United Nations Evaluation Group
UNFPA	United Nations Population Fund
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNIDO	United Nations Industrial Development Organization
UNITAR	United Nations Institute for Training and Research
UNODC	United Nations Office on Drugs and Crime
UNRCO	United Nations Resident Coordinator's Office
UNSDCF	United Nations Sustainable Development Cooperation Framework
USAID	United States Agency for International Development
VNR	Voluntary National Review
VOPE	Voluntary Organization for Professional Evaluation
WFP	World Food Programme

Executive summary

Introduction

The global 2030 Agenda for Sustainable Development recognizes the critical role of evaluation in translating its transformative vision into reality. It calls for review processes to be "rigorous and based on evidence informed by country-led evaluations and data which are high-quality, accessible, timely, reliable, and disaggregated", and highlights the opportunity and need to support and develop national evaluation systems and capacities. In addition, on 19 December 2014, the United Nations General Assembly passed a resolution on "Building capacity for the evaluation of development activities at the country level" (GA69/237).

In 2021, the United Nations Evaluation Group (UNEG) mandated its National Evaluation Capacity Development (NECD) Working Group to prepare a progress report on the implementation of Resolution 69/237. The objective of this report is to record progress on NECD activities by United Nations agencies since the adoption of the 2014 Resolution and suggest areas for improvement. It focuses primarily on support by United Nations agencies to national evaluation systems, with additional reference to other development partners.

Initially, the Working Group developed a working definition and theory of change for NECD. NECD is defined as "the process whereby State and non-state entities and individuals expand, reinforce and sustain national capacity to manage, produce and use evaluation." The theory of change postulates that, in order for governments to effectively develop and implement policies and programmes that positively impact on citizens' lives, and that will help lead towards the achievement of national development goals and the Sustainable Development Goals (SDGs), they need to base their decision-making on timely and credible evidence, including evaluative evidence. To support this process, public institutions need to produce quality evaluations, which in turn requires sufficient capacity at institutional, organizational and individual levels, as well as a conducive or enabling environment, including a culture of evaluation.

The Working Group then commissioned a study of United Nations support to NECD and the evolution of selected national evaluation systems. This included a document review, six country case studies (Benin, Costa Rica, Kenya, Morocco, Philippines and Sri Lanka), a survey of United Nations agencies, and interviews with members of the UNEG NECD Working Group, directors of evaluation offices and key external stakeholders.

Progress on NECD and United Nations agency support

UNEG members have been providing direct NECD support through a wide range of activities targeting governments and non-governmental actors. All countries covered in the case studies received support for the development of evaluation policies and legislation to varying degrees, with the United Nations Development Programme (UNDP) and United Nations Children's Fund (UNICEF) historically being the most engaged. United Nations agencies work with evaluation champions including parliamentarians, voluntary organizations for professional evaluation (VOPEs) and academia, to raise awareness and advocate for evaluation. Support has also been provided through the development of tools such as guidelines, standards and competency frameworks, which are necessary to translate policies and laws into practice. The most common contribution has been to individual capacity development, through trainings, communities of practice, and learning-by-doing through joint and country-led evaluations.

In addition, United Nations agencies increasingly rely on local experts to conduct their own evaluations, and most have been involving governments and non-governmental actors as members of reference groups, steering committees, etc., though with different levels of participation. There are also some examples of joint and country-led evaluations, all of which indirectly contributes to NECD.

Evidence from the country case studies suggests that there has indeed been progress on NECD since the 2014 General Assembly Resolution. The case studies cover some countries with evaluation systems that have emerged since 2014 (Kenya, Philippines), some which were already established (Benin, Costa Rica, Morocco), and one which had advanced in the 2000s but then regressed and has been moving forward again since 2014 (Sri Lanka).

Of the case study countries, Costa Rica (an upper middle-income country) and Benin (classified as a low-income country until 2020) have progressed steadily and have the most functional systems. In Benin, government-led evaluations were initially funded mainly by development partners, but now more than half of the costs are covered by the Government, illustrating that evaluation funding can be catalytic. Morocco has continued to strengthen evaluation through a strong independent-government agency, the National Observatory of Human Development (ONDH), but the system is not government-wide. The Philippines has moved forward, although not as much as might have been expected, since their evaluation policy framework was approved in 2015. Sri Lanka has continued to be active in the VOPE space, and from 2018 has seen some progress in government and parliament. Kenya has also seen some progress with active VOPEs, and approved guidelines in 2019–20. In the latter two cases, government evaluations are yet to be commissioned.

While policies are in place, implementation is still a challenge. Translating evaluation laws and policies into practice involves the development of several elements of a national evaluation system, which need to be built progressively, requiring long-term and consistent support. A key message of the study is the importance of conducting

evaluations to demonstrate their potential, even before all the elements of a system are in place. Useful evaluations can help to generate interest and political support to further develop the evaluation system.

Currently, evaluations are still mostly commissioned by development partners, with a limited number of evaluations produced by governments. This results in limited ownership and use of the results, as well as limited opportunities for the active involvement of government officials in evaluation processes. United Nations evaluations can be used to demonstrate to governments how the evaluation system can be developed. United Nations agencies can also fund government evaluations of national programmes or policies, although only a few agencies are currently doing this.

In addition to direct NECD support, there is potential for United Nations agencies to reconsider the way they conduct evaluations, which could ultimately strengthen the use of evaluations. In order to strengthen national capacity while meeting agency evaluation requirements, United Nations agencies can meaningfully engage government counterparts in evaluation governance mechanisms, consider using national evaluation systems or tools once in place, and promote truly joint or country-led evaluations. United Nations Sustainable Development Cooperation Framework (UNSDCF) evaluations provide great opportunities to use national systems, where capacity is in place. Applying a NECD lens may require the adoption of flexible approaches to make evaluations more responsive to the policy needs and ways of working of country partners.

There are multiple entry points for NECD across United Nations agencies, including: SDG-based policy planning and implementation; the preparation of Voluntary National Reviews (VNRs); support to national development planning and budgetary processes;

National Evaluation Capacities (NEC) Conference 2019, Egypt

©UNDP IEO

advocacy with leadership and parliament; analyses of monitoring and evaluation (M&E) systems; and funding of priority evaluations. Support to NECD also needs to ensure that emerging national evaluation systems are prepared to meet the ever-growing challenges to the achievement of the SDGs. This requires information on what is (and is not) working to combat crises, notably around climate and ecosystem emergencies, and the persistence of high levels of inequality. The pandemic has forced innovation in methods, including the use of rapid evaluation, which need more attention. This is especially relevant given the emerging policy landscape characterized by volatility, uncertainty, complexity and ambiguity. While the counterparts of United Nations agencies are mostly governments, the study also stresses the importance of involving non-state actors in national evaluation systems development, as well as establishing mechanisms to ensure civil society participation in evaluation processes.

United Nations operations can be fragmented, with many barriers to collaboration, and this can weaken, rather than strengthen, national capacity. There is scope to strengthen collaboration and synergy among United Nations agencies and beyond, in particular with the Global Evaluation Initiative (GEI) and EvalPartners. Identifying NECD as a specific programme area, under governance of the UNSDCF, would allow agencies to collaborate further.

Reflections for UNEG

UNEG has 53 diverse members, with mandates ranging from atomic energy to child poverty. Members range from very big agencies with many evaluation staff to very small agencies where 'evaluation' is one person. Agency views on the UNEG role in NECD are as diverse; some consider NECD to be very important, while for others, including those that arguably do not have the resources or country presence for significant investment in NECD, it is a "non-priority".

However, Agenda 2030 and the decade for transformation make specific references to evaluation, pointing to the importance of developing country evaluation capacity. The 2014 General Assembly Resolution reaffirms the importance of national capacity for the evaluation of development activities. The UNEG Norms and Standards now include a Norm on national evaluation capacity.

There is thus a call for the United Nations to invest in NECD across the board, not just through small, individual agency efforts, but collectively, to advance rights, governance, policies and budgets for those left behind. One United Nations respondent highlighted the importance of leadership and determination for this: "the 2014 UN Resolution happened as there was strong leadership by a number of people who worked together in a concerted manner. Without leadership NECD won't happen. This should be at all levels, director, technical."

United Nations agencies should not address evaluation capacity development in isolation, but link NECD to broader evidence generation and use for planning and budget decision-making, which in turn will contribute to better development results.

At the same time, stronger national evaluation systems and capacity will contribute to stronger and more useful agency evaluations. In this light, NECD can be seen as part of all agency mandates. This implies that a minimum part of the evaluation budget should be allocated to NECD. At country level, and in line with United Nations reform to strengthen working together, NECD should appear as a programmatic area in UNSDCFs, under governance, to enable agencies to come together.

There is room to create a shared framework for NECD work, and the possibility to develop common work plans. While there is not yet a shared language and understanding in UNEG of how to promote NECD, more can be done, even with limited resources. The 2012 UNEG Practical Tips Guideline (UNEG, 2012) provides a good basis for this, and could be enhanced with the experience gained over the last ten years. There could also be a minimum standard for United Nations agencies to support country-led SDG-related evaluations.

UNEG could provide resources on NECD, potentially in collaboration with other partners such as the GEI, for example: developing a public repository of available NECD resource materials and tools; distilling lessons on what works in NECD; or documenting and sharing stories of successful NECD initiatives. These resources and tools could be shared with United Nations country teams and NECD focal points and used for advocacy, conferences and discussions with different partners outside of the evaluation community. In the words of one respondent:

> **"Ultimately, evaluation in the United Nations should include the vision we have for our partner countries: national evaluation being carried out of national programmes and feeding back into policy and practice related to the SDGs. These changes are possible."**

Recommendations

With a view to contributing to effective sustainable development and strengthening in-country evaluation capacity, United Nations agencies and their evaluation functions should implement the following recommendations, in line with their capacity and contexts.

Recommendation 1. All United Nations agencies should conduct their evaluations in a way that fosters national capacity development.

1.1 In principle, all United Nations agency country programme evaluations and UNSDCF evaluations should include the meaningful presence of national governments in management structures (reference groups, steering committees), with countries playing a leading role in such governance mechanisms. The next update of the UNSDCF evaluation guidelines should place stronger emphasis on supporting NECD and incorporate these recommendations into the text.

1.2 When feasible, United Nations agencies should foster joint and country-led evaluations. United Nations entities should also support country-led government

evaluations that address agency priorities, which may then count towards United Nations agency evaluation coverage requirements.

1.3 In countries with national evaluation systems, United Nations agencies should consider using national evaluation plans, guidelines, standards and other relevant elements in the conduct of their own evaluations, so as to respect the countries they are working in and enhance the credibility of these systems. In the process, this may help to further strengthen the national evaluation system.

1.4 United Nations agencies should commit to increase the numbers, and strengthen the capacity, of local evaluators, including through support to young emerging evaluators within evaluation teams. Agencies should consider using evaluation consultants as facilitators to build evaluation capacity, in addition to their standard evaluation responsibilities.

Recommendation 2. In line with General Assembly Resolution 69/237, United Nations agencies and their evaluation functions should continue to support the capacity development of national evaluation ecosystems, including support to the enabling environment, institutional and individual capacity.[1] This may include a range of actions, aligned with the context and demand, as proposed in the following sub-recommendations:

2.1 United Nations agencies should, in collaboration with relevant partners and stakeholders, support country-owned M&E systems analysis to identify strengths and weaknesses in the broader ecosystem, followed by support to the definition and implementation of a medium-term evaluation capacity development strategy, according to their comparative advantages.

2.2 United Nations agencies should support the engagement of senior policymakers in the executive and parliament to increase their exposure to evidence-informed policy and practice, and respond to their evidence needs. United Nations agencies should support the development of a policy and regulatory environment to enable and sustain useful and credible evaluation processes and practices; as well as the strengthening of institutional capacity, frameworks and processes for conducting and using evaluations. Support for specific country-led or joint evaluations should be considered to pilot-test instruments, promote opportunities for learning-by-doing, and demonstrate the usefulness of evaluation.

2.3 United Nations agencies should facilitate the engagement of non-state actors in the evaluation ecosystem, including VOPEs, academic and training institutions, citizens able to engage with evidence and policymaking debates, as well as evaluation professionals.

1 The evaluation "ecosystem" is broader than the government system, including other systems and players that may contribute to the practice of evaluation in a country, such as parliaments, universities and VOPEs.

2.4 United Nations agencies should advocate for the integration of the SDGs, principles of gender equality, human rights, leave no one behind and disability inclusion, and climate change issues in country-led evaluations and national evaluation systems.

Recommendation 3. All United Nations agencies should coordinate and collaborate on NECD at corporate, regional and country levels, allocating adequate time and resources.

3.1 United Nations agencies should explicitly include NECD as part of their mandates, incorporated into their evaluation policies, and allocate time and resources at corporate, regional and country levels. At least 10 percent of evaluation resources should be allocated to NECD.

3.2 United Nations agencies should ensure inter-agency information sharing, coordination and collaboration on NECD at corporate, regional and country levels.

3.3 At country level, United Nations agencies should include NECD as an explicit part of individual agency country programmes and of the UNSDCF, for example under a governance outcome to be monitored by the United Nations country team M&E Working Group. When more than one agency is supporting NECD in the same country, they should coordinate efforts under a joint NECD programme, managed through a country-led steering group chaired by key government M&E champions and involving all actors in the evaluation ecosystem.

1. Introduction

1.1 Introduction

The world is facing multiple crises, from the current COVID-19 pandemic, to ever more present signs of climate and ecosystem breakdown, and the dangers of social disruption arising from persistent inequalities. The United Nations is all too aware of this, being at the front end of humanitarian action and development cooperation. The Sustainable Development Goals (SDGs) were adopted in September 2015 with the aim of addressing these systemic challenges, and the global 2030 Agenda for Sustainable Development proposes a transformative vision to address multiple crises.

Agenda 2030 recognizes the critical role of evaluation in translating its vision into reality, calling for review processes to be "rigorous and based on evidence informed by country-led evaluations and data which are high-quality, accessible, timely, reliable, and disaggregated." To this end, it highlights the opportunity and need to support and develop national evaluation systems and capacity, while enhancing alignment and learning across different national contexts, to the most practicable extent possible (Librado & Maclean, 2019). Meanwhile, on 19 December 2014, United Nations Resolution GA69/237 was passed to promote the development of country-led evaluation and national evaluation capacity (NEC).

Evaluation as we currently know it was brought to the Global South in the 1950s by the international development community, including the United Nations system. However, development partners initially focused only on evaluation of their own interventions (ILO, 2009). Attention to national evaluation capacity development (NECD) emerged in the 1990s, including from United Nations agencies. Colombia established its evaluation system in 1994 and Sri Lanka during the 2000s. In 1999, the United Nations Children's Fund (UNICEF) supported national monitoring and evaluation (M&E) associations, known as voluntary organizations for professional evaluation (VOPEs) from Comoros, Kenya, Madagascar, Malawi, Niger and Rwanda to establish the African Evaluation

Association (AfrEA) (ILO, 2009). The first national evaluation agencies were created in Colombia and Costa Rica in 1994, in Sri Lanka around 2002, in Mexico in 2005 (CONEVAL), in Benin in 2007 (Bureau of Evaluation of Public Policy and Analysis of Government Action or BEPPAAG), followed by other African countries. In 2009, the United Nations Development Programme (UNDP) convened its first biannual National Evaluation Capacities conference, bringing together 55 representatives of national governments, VOPEs, academia and United Nations agencies from 30 countries (UNDP IEO 2019). UNICEF played a crucial role in launching the EvalPartners global initiative in 2012, which aimed to promote awareness of the role of effective evaluation in decision-making among policymakers, public opinion and other key stakeholders.

In 2012, the United Nations Evaluation Group (UNEG) Task Force on NECD released some practical tips for United Nations staff to strengthen national evaluation capacity systems (UNEG 2012). The report established a new paradigm for evaluation capacity development (ECD) support within the context of developing a country-owned national evaluation system, highlighting that the driver for NECD should be good governance rather than donor evaluation and reporting requirements. The report also pointed out that, since national M&E systems were at different stages of maturity, United Nations NECD support should be context specific.

Adoption of the new NECD paradigm centred around national ownership and linking evaluation capacity to a national vision, accountability and good governance. This laid the foundations for the adoption of the United Nations General Assembly Resolution A/RES/69/237 "Building capacity for the evaluation of development activities at the country level" in December 2014.

A new United Nations Resolution is now being considered to build on that of 2014. The objective of this report is to record progress in NECD activities by United Nations agencies since the adoption of the 2014 Resolution and suggest areas for improvement. The report focuses primarily on support by UNEG member agencies. It also refers to other initiatives (including by Member States) where relevant, including a set of country case studies conducted as part of this study.

The report is based on background papers commissioned by the UNEG NECD Working Group, which was established in 2019 (members of the Working Group are found in Appendix 3). The report is addressed primarily to UNEG and United Nations agencies.

1.2 Structure of the report

The report structure is as follows**:**

- **Section 1** introduces the report, and presents the scope and methodology employed.
- **Section 2** provides the working definitions used by UNEG with some commentary, and describes the theory of change that underlies it, mapping some key elements.

- **Section 3** describes United Nations agency support to NECD and progress in the development of national evaluation systems, drawing from the case studies, interviews with key respondents, the survey and literature review.
- **Section 4** draws out lessons and reflections for UNEG going forward.
- **Section 5** provides conclusions and recommendations.

Various appendices share background information on the methodology and key inputs:

- **Appendix 1.** Extracts from relevant United Nations resolutions.
- **Appendix 2.** Survey of UNEG Members on National Evaluation Capacity Development.
- **Appendix 3.** Members of Working Group.
- **Appendix 4.** People interviewed outside the Working Group.

1.3 Towards a new United Nations General Assembly Resolution

1.2.1 Timeline and key milestones

This section provides a record of the timeline of significant events related to NECD and the 2014 Resolution, leading up to a potential new Resolution in 2022. Starting with previous Resolutions which informed the 2014 Resolution (see Appendix 1 for extracts), it makes reference to other milestones. The key elements are:

- **Resolution 59/250 of 22 December 2004**. This asked UNEG to make further progress in system-wide collaboration on evaluation, in particular the harmonization and simplification of evaluation methodologies, norms, standards and cycles, and strongly encouraged country-level evaluations of United Nations Development Assistance Frameworks (UNDAFs). This Resolution emphasized that national governments have primary responsibility for coordinating external assistance, including from the United Nations system, and evaluating the impact of its contribution to national priorities.
- **The Paris Declaration of 2005 and Accra Agenda for Action**. These focus on "Ownership: Developing countries setting their own strategies for poverty reduction, improve their institutions and tackle corruption," as well as "Alignment: Donor countries should align behind these objectives and use local systems." This is at the root of the NECD focus (see **Box 1**).[2]
- **United Nations agency evaluation policies.** UNDP adopted its first evaluation policy in 2006, which references Resolution 59/250, and under the guiding principle of national ownership mentions that "it should build the capacity of national institutions to implement, monitor and evaluate." Since then, other United Nations agency evaluation policies have also referenced NECD. UNICEF followed in 2007, and its current policy discusses NECD extensively with responsibilities for staff at headquarters, regional

2 https://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm

Inception meeting with Governmet of Eswatini during WFP joint evaluation of the school feeding programme

and country office levels. The United Nations Population Fund (UNFPA), UN Women and the World Food Programme (WFP) also cover NECD in their policies, and the International Labour Organization (ILO) focuses on its constituents. However, other agencies, including the United Nations High Commissioner for Refugees (UNHCR), United Nations Office on Drugs and Crime (UNODC), and the Food and Agriculture Organization (FAO), include little or no mention of NECD or country-led evaluation in their policies.

- **Resolution 62/208 of 19 December 2007**. Article 10 of this Resolution requests the United Nations Development System to pursue the full integration of operational activities for development at country level, with national planning and programming under the leadership of national governments at all stages of the process, while ensuring the full involvement of all relevant stakeholders at the national level. Article 12 emphasizes that programme countries should have greater ownership and leadership in the evaluation of all forms of assistance, including that provided by the United Nations Development System, and requests the United Nations to pursue and intensify its efforts to strengthen evaluation capacity in programme countries.[3]
- **Resolution 67/226 of 21 December 2012**. Article 61 of this Resolution calls upon the United Nations Development System to strengthen its focus on developing national capacity for: development planning; disaggregated data collection and analysis; implementation; reporting; and M&E. Section G of Article 175 emphasizes that programme countries should have greater ownership and leadership of the evaluation of the assistance provided by the United Nations Development System. It calls upon members of the United Nations Development System to intensify efforts

3 Meanwhile in 2012, EvalPartners was formed by the International Organization for Cooperation in Evaluation (IOCE) and the UN as an innovative partnership.

to strengthen national evaluation capacity in programme countries and develop and implement guidelines for further strengthening national evaluation capacity for development operational activities. There is also a specific section on system-wide evaluation.

- **Resolution GA69/237 of December 2014.** This is a very short document, and the first stand-alone Resolution on building capacity for the evaluation of development activities at country level. This Resolution confirms the importance of building national capacity for the evaluation of development activities and invites United Nations agencies, with the collaboration of national and international stakeholders, to support efforts to further strengthen the evaluation capacity of Member States in accordance with their national policies and priorities. The main elements are shown in **Box 1**.

Box 1. Resolution GA69/237 of December 2014 on building capacity for the evaluation of development activities at country level

Reiterates the importance of building national capacity for the evaluation of development activities.

Reaffirms that national capacity for the evaluation of development activities may be further strengthened by entities of the United Nations Development System upon request and in accordance with the principle of national ownership and with the national policies and priorities defined by Member States.

1. **Notes that evaluation at the country level should be voluntary and carried out upon request by Member States.**
2. **Invites the entities of the United Nations Development System, with the collaboration of national and international stakeholders to support, upon request, efforts to further strengthen the capacity of Member States for evaluation, upon request, and in accordance with their national policies and priorities.**

1.2.2 What has happened since 2014

In 2015, the 2030 Agenda for Sustainable Development was approved, and within the section on follow-up (United Nations, 2015: 74) it was stated that: Follow-up and review processes at all levels will be guided by the following principles:

g. They will be rigorous and based on evidence, informed by country-led evaluations and data which is high quality, accessible, timely, reliable and disaggregated.

h. They will require enhanced capacity development support for developing countries, including the strengthening of national data systems and evaluation programmes, particularly in African countries, Least Developed Countries, Small Island Developing States, landlocked developing countries and middle-income countries.

This explicitly positions country-led evaluation as an integral part of the follow-up and review processes for the SDGs.[4] In essence, this provision can be seen as the policy applicable to all national evaluation systems.

Reporting against the SDGs is done through the Voluntary National Review (VNR) process, which provides a potential channel to strengthen the use of evaluation. However, so far, the use of evaluation in VNRs has been limited. To promote the implementation of this provision of Agenda 2030, UNICEF, UNDP and UN Women supported the production and dissemination of a Briefing Series by EvalSDGs and the International Institute for Environment and Development that advocates for the importance of embedding evaluation in national review processes in the framework of Agenda 2030.

2015 was declared the International Year of Evaluation, and in November 2015 the Global Evaluation Forum brought together key stakeholders to finalize the "Global Evaluation Agenda 2016-2020" (referred to as EvalAgenda 2016-2020) and develop action plans to implement it (see **Box 2**).

Box 2. EvalAgenda 2016-2020

This declaration recognizes that evaluation has enormous potential to help improve society. By influencing policy makers, other key stakeholders and public opinion, evaluation can help to ensure that public policies, programmes and processes are informed by sound evidence and lead to effective and equitable results, thus improving the lives of all people and ensuring sustainable development to protect our planet. However, it notes that evaluation has not yet reached its full potential. The declaration states:

> **"We know that evaluation is not simply a value-neutral management tool. We and EvalPartners are united by a shared commitment to promoting and supporting equitable and sustainable human and ecological development. We promote evaluation processes and criteria grounded in values of equity, gender equality, and social justice and on shared principles of partnership, innovation, inclusivity, human rights and the protection of the planet.**
>
> **We declare our support for evaluation priorities during the first five years of the 15-year period addressed by the new Sustainable Development Goals (SDGs). We understand that the "No one left behind" principle stated in the SDGs is embedded as a key value that goes across the three building blocks of an effective evaluation system – enabling environment, institutional capacities and individual capacities."**

4 United Nations General Assembly A/Res/70/1. Transforming our world: the 2030 Agenda for Sustainable Development. Para 74g.

In 2016, the UNEG Norms and Standards for Evaluation (originally adopted in 2005) were revised to include a new norm on national evaluation capacities (Norm 9). This states that: "The effective use of evaluation can make valuable contributions to accountability and learning and thereby justify actions to strengthen national evaluation capacities. In line with the General Assembly Resolution A/RES/69/237 on building capacity for the evaluation of development activities at the country level, national evaluation capacities should be supported upon the request of Member States."[5] The UNEG Evaluation Competency Framework[6] was also updated, and though the competencies do not specifically refer to NECD, they do include reference to:

Integrating evaluation into policy and programming

- [Officer] Is able to communicate to stakeholders the value of evaluation as a vital component of policy and programming in the achievement of the SDGs.
- [Officer] Is able to support stakeholders in integrating evaluation into policy and programme development and management through knowledge of learning organizations.

Using utilization-focused approaches

- [Officer] Is able to consistently promote the engagement of users and beneficiaries in evaluation processes in order to promote evaluative thinking and the wide use of evaluation findings.

In 2016, the Quadrennial Comprehensive Policy Review mentioned national evaluation capacity in Article 123, stating: 'In order to achieve results in this regard, greater attention to the development of national evaluation capacities will be required. The expectations of the programme countries for the support of the United Nations Development System for genuine strengthening of their national evaluation capacities are not being met … There is need for more innovative and effective methods to establish and improve national evaluation policies, systems and programmes which can inspire country ownership and create the motivation to design and manage country led evaluations, as well as for support from across the development system to enable countries to effectively use the results from such evaluations in their national decision making.' Article 247 states: "Strengthening national results-based management systems requires the entities of the United Nations Development System to make more use of existing national evaluation capacities and monitoring and reporting systems, to support the strengthening of national statistical systems and in order to improve data availability and data quality."

In September 2019, the United Nations Decade of Action to achieve the SDGs was launched.[7] In the same year, UNEG created the NECD Working Group which

5 http://www.unevaluation.org/document/detail/1914 [Cited 29 July 2021]

6 http://www.unevaluation.org/document/detail/1915 [Cited 29 July 2021]

7 https://www.un.org/sustainabledevelopment/decade-of-action/

commissioned this study, UNICEF undertook a study on NECD and the SDGs in Asia-Pacific (from which this study has drawn), and the latest UNEG Strategy 2020-2024 was published. The Strategy is very explicit about NECD, including Strategic Objective 3 to "Influence policy-making and operational work through evaluations" (UNEG, 2019). It later states that UNEG should:

> **"Build on its capacities as well as those of its partners. New methodological approaches will require updating the skill set of UNEG Member staff, as well as a more robust decentralized evaluation capacity. Building national evaluation capacities should be an integral part of UNEG's evaluation efforts, including advocating for greater use of evaluation evidence in Voluntary National Reviews".**

In summary, a series of United Nations Resolutions and declarations, as well as the evaluation policies of United Nations agencies, have called for NECD. Agenda 2030 creates an opportunity and a need to support and develop national evaluation systems and capacity while enhancing, to the extent practicable, alignment and learning across different national contexts.

1.4 Methodology

The UNEG Working Group was mandated by UNEG to prepare a progress report on the implementation of United Nations General Assembly Resolution 69/237. A preliminary report was commissioned by the UNICEF Evaluation Office and prepared by Natalia Kosheleva, following which the Working Group elaborated an outline for a more comprehensive report. The Working Group developed a working definition and theory of change for NECD. The theory of change was developed through an iterative process, working from a proposed desired situation whereby public institutions use evaluative evidence to inform the development and implementation of policies and programmes that positively impact on citizen's lives, helping to lead towards the achievement of national development goals and the SDGs. The group then identified conditions that need to be in place for this to happen, and the types of inputs or activities that would support the creation of these conditions. The theory of change was revised during the preparation of the present report as new ideas emerged throughout the process, and is presented in section 2.2 below.

Ian Goldman was commissioned in April 2021 to help the Working Group develop the report.[8] His support included: expand the literature review and refine the conceptual framework; develop tools for and conduct additional data collection on results and lessons learned from a sample of NECD initiatives and United Nations staff; design a

8 Advisor, Evaluation and Evidence Systems, Center for Learning on Evaluation and Results, Anglophone Africa (CLEAR-AA), University of Witwatersrand; Adjunct Professor, Nelson Mandela School of Public Governance, University of Cape Town; Visiting Professor, University of Reading, UK.

country case study approach; and, prepare a full draft report incorporating elements of the initial report by Kosheleva, and the theory of change and definitions developed by the Working group, as well as other inputs from them. Ian Goldman also carried out case studies on the Philippines,[9] Sri Lanka,[10] and Kenya. Miguel Angel Lombardo Chico prepared case studies on Costa Rica and Morocco, and Emmanuel David-Gnahoui conducted the case study on Benin.

Data collection and analysis was done using a mixed methods approach, though predominantly qualitative. The key elements were:

- **A document review**, including data on United Nations Resolutions, United Nations and other reports, reports on NECD, data on UNEG NECD initiatives compiled by the Working Group, and a review of evaluation policies of some key agencies.
- **A survey** of United Nations agencies to explore what they were doing in NECD, using the UNEG NECD theory of change as a basis (discussed in section 2). The survey used some closed questions for aggregation purposes, but also open questions to explore methods and lessons. The survey was sent by the UNEG Secretariat and 14 agencies responded, many having consulted internally in regions and countries. Responding agencies included: FAO, the Global Environment Facility (GEF), UN-Habitat, ILO, International Organization for Migration (IOM), United Nations Office of Internal Oversight Services (OIOS), Pan American Health Organization (PAHO), UNDP, UNFPA, UNHCR, UNICEF, United Nations Industrial Development Organization (UNIDO), United Nations Institute for Training and Research (UNITAR), UNODC, UN Women and WFP.
- **Interviews** with various stakeholders, including: 15 Members of the Working Group to deepen the insights that emerged from the survey (FAO, Habitat, OIOS, PAHO, UNDP, UNFPA, UNICEF, UNODC, UN Women, and WFP); four directors of evaluation (FAO,[11] UNDP, UNFPA, WFP); and five stakeholders with an understanding of the external environment and roles of other agencies in the NECD space. The full list of interviewees is in Appendix 4.
- Six **case studies** to understand the dynamics of NECD and the roles played by United Nations agencies.
- A **workshop** with the Working Group on the main findings and recommendations in September 2021.

The case studies were purposively selected to include countries where, ideally, at least two United Nations agencies had been involved in NECD, and some elements of national evaluation systems were in place. It was decided not to include some of the well-known and documented examples such as Colombia, Mexico and South Africa although

9 Complementing a UNICEF case study completed in 2019 (CPRM Consultants, (2019).

10 Complementing a UNICEF case study completed in 2018 (Trikawalagoda, 2018).

11 Also the UNEG chair.

these are referenced in the report.[12] The eventual selection (Benin, Costa Rica, Kenya, Morocco, the Philippines and Sri Lanka) covers a range of characteristics, as follows: [13]

- They either have functioning evaluation systems through which evaluations are undertaken (Benin, Costa Rica, Morocco), or less strong national evaluation systems where some elements such as a policy are in place, but which are not yet systematically implementing evaluations (Kenya, Philippines, Sri Lanka);
- They include one low-income country (Benin)[14] while the rest are middle-income countries; and
- They cover multiple regions, including Latin America (Costa Rica), Francophone Africa (Benin), Middle East and North Africa (Morocco), Anglophone Africa (Kenya), and South and East Asia (Sri Lanka, Philippines).[15]

Each case used standard research instruments, including:

- **Interviews** with over ten respondents per country drawn from government, VOPEs, United Nations agencies and in some cases the United Nations Resident Coordinator's Office (UNRCO) (a total of 76 interviewees);
- **Desk review** of substantial documentation;
- **Search** of the main evaluation repositories to see how many of the evaluations done in each country could be accessed publicly.

The report structure and instruments were guided by the UNEG theory of change (see section 2).

Limitations of the study include: the absence of baseline data for 2014/15; the moderate response rate to the survey by UNEG members and thus potential gaps in analysis of United Nations agency response to the Resolution; and limited generalizability of progress in national evaluation systems, given the limited number of country case studies. The latter, however, were complemented by information from other countries and studies through the literature review. The Working Group is confident that the analysis provides a firm basis for the development of future strategies for UNEG support to NECD.

The final report and recommendations reflect multiple consultations with, and intense engagement of, members of the UNEG NECD Working Group to arrive at the final text.

12 Note as one of the authors led the South African national evaluation system he has very good data from that country, which is sometimes used as comparison.

13 In the case of the Philippines and Sri Lanka, case studies had been done in 2018–2019 through a UNDP/UNICEF project looking at countries in the Asian region. These case studies were drawn from extensively to not re-invent the wheel and were updated and adapted for the purposes of this study.

14 Benin is classified since 2020 as a lower middle-income country, but much of its national evaluation systems development occurred when still considered a low-income country.

15 Sri Lanka and Philippines were also selected as they had existing UNICEF-funded case studies which could be built on (Trikawalagoda, 2018; CPRM, 2019).

2. National evaluation capacity development

2.1 Definitions and concepts

The UNEG NECD Working Group defines NECD as "the process whereby state and non-state entities and individuals expand, reinforce and sustain national capacity to manage, produce and use evaluation." **Box 3** presents the full Working Group definition.

Box 3. UNEG Working Group definition of NECD

National Evaluation Capacity Development is the process whereby state and non-state entities and individuals expand, reinforce and sustain national capacity to manage, produce and use evaluation.

NECD is linked to national priorities and ultimately aims at strengthening governance through accountability and learning, thereby improving development and peace outcomes, in terms of country priorities and Agenda 2030, contributing to human rights and equity.

Greater demand for and better quality of evaluations, and their use in policy and practice require individual skills and knowledge, institutional systems and policies, and an enabling environment, including a conducive evaluation culture.

NECD supports, through direct or indirect efforts, the strengthening of a country-owned national system within a particular cultural, social and political context.

While NECD is conducted primarily at national level, it can also build strong linkage and synergies at subnational, regional and global levels.

The Working Group also developed a set of definitions and approaches which are used in this report (see **Table 1**).

Table 1. Definitions linked to NECD

Term	Explanation
Evaluation	An evaluation is an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using appropriate criteria such as relevance, effectiveness, efficiency, impact and sustainability. An evaluation should provide credible, useful evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision-making processes of organizations and stakeholders. (UNEG, 2016)
Monitoring	Management's continuous examination of any progress achieved during the implementation of an undertaking in order to track its compliance with the plan and to take necessary decisions to improve performance. (UNEG, 2016)
National evaluation system	National evaluation systems are specific to each country but are influenced by the global environment. They can include several components such as: designated public institution(s) for evaluation, an evaluation policy, national evaluation guidelines, standards or ethics.
National evaluation ecosystem	The term "National evaluation system" is often used to refer to the government evaluation system, but there is a broader evaluation "ecosystem", which includes other systems and players that may contribute to the practice of evaluation in a country, such as parliaments, universities and VOPEs. Parliaments play an important role in evaluation in some countries. Many developing countries already have national VOPEs that bring together individuals and organizations interested in evaluation. VOPEs provide a platform for peer exchange among evaluation professionals, and contribute towards developing capacity on evaluation supply and demand through training. In some cases, they are able to play a catalytic role by advocating for national evaluation systems and providing technical support to governments in the process of establishing or operating national evaluation systems, as well as supporting academia to launch courses on evaluation. Donors and development partners with an explicit NECD mandate also play an important role in the national evaluation ecosystem by supporting government, VOPE and academia efforts to develop national evaluation systems. In addition, donors and development partners play a role as commissioners and users of evaluation.

Table 1. Definitions linked to NECD

Term	Explanation
ECD versus NECD	NECD differs from ECD in the following: • NECD targets national development / humanitarian / peacebuilding actors; • For United Nations agencies, NECD aims to contribute to the outcomes of Agenda 2030; • NECD focuses on building national evaluation systems applied in a particular cultural, social and political context and on progress towards good governance.
Developing versus building	NECD is about developing and strengthening capacity in the long-term. Capacity 'building' suggests starting from scratch, whereas capacity development is a long-term effort that needs to be embedded into broader change processes owned and driven by those involved. NECD is context-specific and about changing values and mindsets as well as acquiring new skills and knowledge. However, in Agenda 2030 these terms are used interchangeably.
Scope and types of NECD interventions	NECD works at three levels: individual capacity, institutional capacity and the enabling environment. The range of interventions that contribute to NECD can be broad and vary according to contexts. Different country contexts have different needs and, in some contexts, institutional weaknesses or political instability may require more work with non-governmental actors. NECD support, regardless of the type of activity, will focus on supporting progress at national, subnational, or sectoral/line ministry level.
Direct and indirect support to NECD	NECD, particularly in the context of UNEG, includes both targeted direct actions and indirect interventions that relate to how the United Nations agencies carry out their evaluation and related work. For example, including local actors and finding ways to expose them to training and awareness-raising.
Actors	For NECD, these include government departments, civil society organizations (CSOs), academia, VOPEs, research centres, parliamentarians, evaluators (including young evaluators), etc.

Source: UNEG NECD Working Group unless other sources cited.

NECD is seen as a driver for the establishment of effective national evaluation systems, which necessarily include the following characteristics or enabling conditions:

- National leadership and key decision makers understand the value and potential contribution of evaluation to the achievement of national goals, and are committed to establish a national evaluation system and sustain it over a long period of time.
- A strong civil society, including VOPEs, demanding and advocating for evidence-based policymaking and a strong national evaluation system.

- Infrastructure to ensure systematic, comprehensive and credible approaches to evaluation, including national evaluation policies and standards, technical evaluation guides, as well as documents that guide ongoing NECD efforts.
- Availability of skilled people to commission and conduct evaluations.
- Capacity within government institutions to incorporate and use evaluation findings as part of the normal process of business.
- Capacity to support ongoing NECD efforts, for example through training and technical advice.

It should be noted that there are different ways to categorize elements of national evaluation systems. For example, Goldman et al. (2019) adapt Holvoet's definition as shown in **Box 4**.

Box 4. Holvoet's Six Descriptive Characteristics of a National Evaluation System

Holvoet characteristic	Elements
Policy	Evaluation plan; approach to monitoring vs. evaluation; autonomy and impartiality; feedback to management; alignment to planning and budgeting.
Methodology	Selection of results and areas to be evaluated; priority setting; evaluation methodologies used; data collection and quality.
Organization	Coordination and oversight; statistical office; line ministries; decentralized levels; link with interventions.
Capacity	Problem acknowledged; capacity development plan.
Participation of wider stakeholders	Parliament; civil society (including VOPEs); universities; donors; private sector.
Use	Effective use of evaluation; internal usage of evaluation findings.

Source: Goldman *et al.* (2019), adapted from Holvoet and Renard 2007.

National governments in United Nations programme countries have demonstrated a growing interest in evaluation. At the same time, recent research indicates that the conditions necessary for an effective national evaluation system are still unmet in many countries. For example, a study of the evaluation landscape in Africa conducted by the Centre for Learning on Evaluation and Results (CLEAR) Anglophone Africa and the Centre for Research on Evaluation, Science and Technology found that:

- Evaluators in Africa are largely "parachuted" in from the North. A disproportionate number of evaluators from the global North are leading evaluation teams and research initiatives around evaluations in Africa.

UNODC workshop: "Independent Evaluation and National Evaluation Capacity Development" with the Master in Evaluation of the University Moulay Ismail of Meknes, Morocco.

- M&E in the region is overwhelmingly dominated by monitoring. Even evaluation practice focuses heavily on whether sufficient progress is being made towards predetermined results, and less on overall programme effectiveness and strategic planning.
- The quality of evaluations varies significantly (EvalForward 2020).

2.2 A theory of change for NECD

The UNEG NECD Working Group theory of change to support NECD was developed during the process leading up to this research, and revised during preparation of this report. As the case studies were carried out and analysis deepened, the theory of change was adapted in response to new reflections.

Given that, in different contexts with different levels of national evaluation system development, pathways and related theories of change will be different, the theory of change presented here is necessarily a synthesis. It also focuses on elements that are easier to influence, such as institutional frameworks and capacity. The enabling environment for the use of evaluative evidence for decision-making is a function of a broader political economy which is not easily influenced through typical development interventions.

The UNEG theory of change for NECD (**Figure 1**) postulates that, in order for governments to effectively develop and implement policies and programmes that positively impact on citizens' lives, and that will help lead towards the achievement of national development goals and the SDGs, they need to base their decision-making on timely and credible evidence. There are many types of evidence that may be used. Statistics produced by national bodies can inform decision-makers of progress towards national or local development targets and the SDGs. Monitoring and reporting systems can inform on progress in implementing programmes on the ground. Good quality evaluations provide information on whether programmes and policies are having the desired effect or not, how, for whom, under what circumstances and why, and how these can be strengthened.

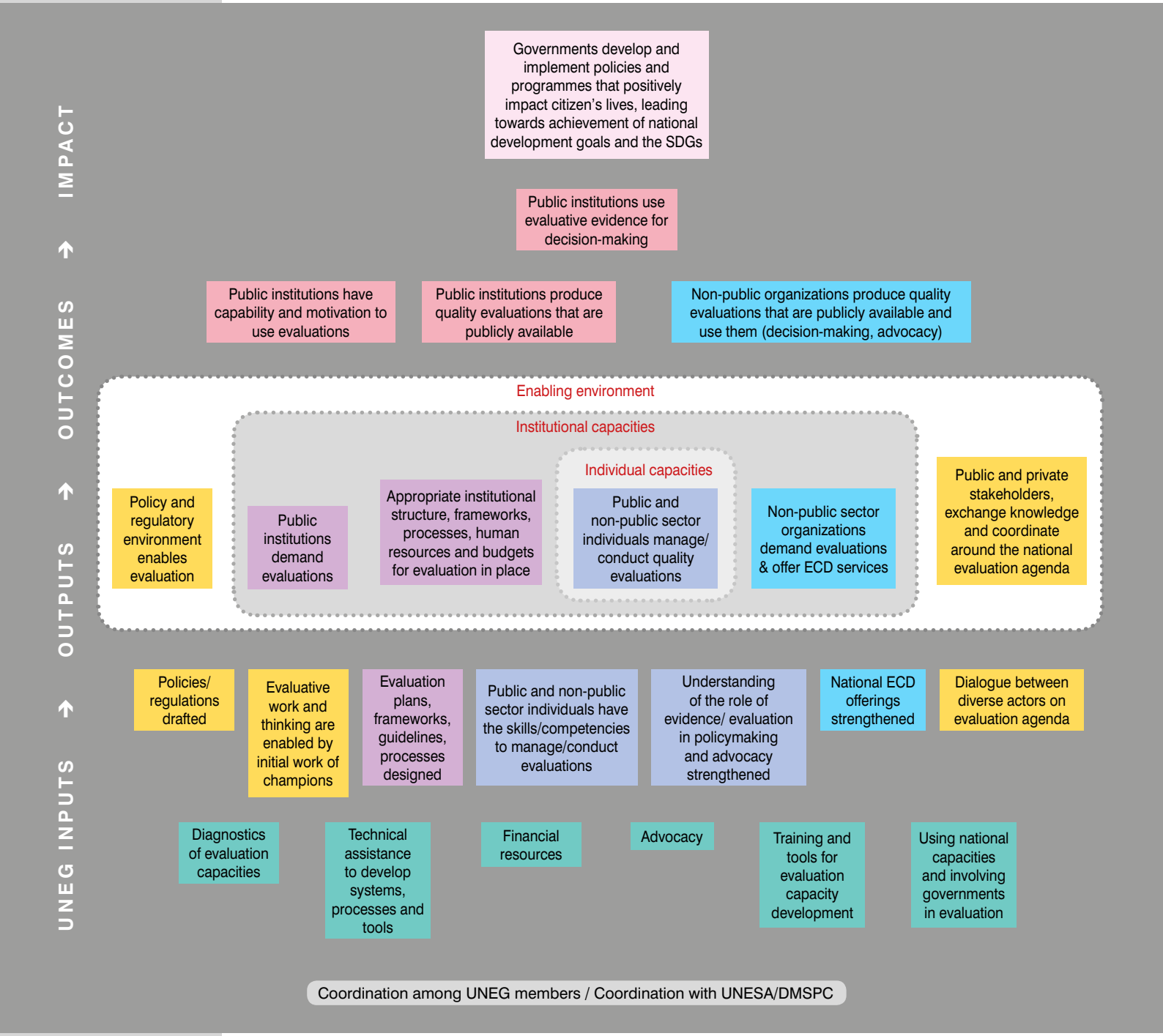
To support this, public institutions need to produce quality evaluations, which need to be publicly available, not only to policymakers but also other stakeholders engaged in advocacy and policy dialogues. In addition, non-public institutions can and should support this process by producing and sharing complementary evaluations, as well as playing appropriate roles in the wider national evaluation ecosystem.

For public (and non-public) institutions to produce quality and useful evaluations (i.e. that meet predefined norms for evaluation coverage and quality standards, are conducted in an ethical manner, and integrate gender, human rights and the principle of leaving no one behind), there is a need for sufficient capacity at institutional or organizational, as well as individual, levels. Furthermore, wider environmental factors including policies, frameworks and available budgets influence the capacity of institutions to produce evaluations and use them to help achieve national development objectives.

An enabling environment for evaluation requires the political will or motivation of key actors within the executive and legislative bodies to produce and use evaluations, which may be influenced through targeted advocacy efforts. However, the broader governance environment is equally important, as political systems and political cultures influence access to information, the strength of civil society, openness to critique, accountability mechanisms, etc. Influence at this level may be beyond the scope of evaluation capacity development interventions but needs to be considered as part of identifying and monitoring risks and assumptions. More specifically, a strong policy and regulatory environment is important to enable and sustain evaluation processes, practices and use, which may initially arise due to the work of a few "champions".

At institutional level, for example, within an entity designated to coordinate the national evaluation function, or entities responsible for monitoring and evaluation in line ministries, there is a need for frameworks and processes, as well as adequate human resources (staffing) and budgets for evaluation. Demand for evaluation must exist, for example from legislators, parliamentarians and/or policymakers within ministries (this can also be mirrored at the subnational level). Demand can emanate from civil society as well, and ideally, public and private stakeholders also need to exchange knowledge and coordinate around the national evaluation agenda as part of the overall enabling environment.

Figure 1. UNEG theory of change for national evaluation capacity development



A strong national evaluation ecosystem will also benefit from institutional capacity in non-public sector entities which could advocate for, commission and/or conduct evaluations. This includes VOPEs, academic institutions, private companies, think tanks and others.

UNDP IEO

National Evaluation Capacities (NEC) Conference 2015, Thailand

At individual level, the public sector needs a critical mass of staff with the knowledge, skills and competencies to develop and implement appropriate evaluation plans, commission and manage evaluations and, depending on the context, conduct evaluations. Outside of the public sector, there is a need for qualified evaluators to conduct evaluations.

Thus there are several entry points for ECD services, including from United Nations agencies, all of which will depend on the country-specific context and level of existing capacity. As a first step, advocacy to strengthen understanding of the role of evidence and evaluation in policymaking may be appropriate, followed by support to carry out assessments of evaluation capacity and needs and the development of ECD strategies.[16] Technical assistance can be provided to help draft policies or legislation, drawing on good practices from other countries or related contexts (including through South-South cooperation, peer exchange, etc.), to help institutions develop the necessary systems, processes and tools to identify priorities for, plan, commission or conduct evaluations, and use the results in decision-making processes. Experience suggests that it can be effective to support, from an early stage, evaluations that

16 Assessments can cover all three levels – the enabling environment, institutional and individual capacity.

serve as prototypes for how the system might work, while also advocating for further evaluations.

Training can be provided to develop the capacity of individuals to manage and conduct evaluations. Support can be given to national institutions to strengthen their own ECD offerings, which will be important to ensure that individual capacity continues to be built (so that, for example, the transfer or retirement of previously trained and experienced staff in public institutions does not undermine the sustainability of the evaluation ecosystem). This may include training of trainers, assistance to develop university curricula on evaluation, and support to national institutions providing civil service training to develop and offer courses on evaluation and the use of evidence for informed decision-making.

UNEG member agencies (and other international development partners) can indirectly support NECD by using national capacity for the evaluations they commission. At the individual level, agencies can prioritize hiring evaluators from the country or region. At institutional level, agencies can liaise with relevant national counterpart institutions and, where possible, conduct evaluations jointly with the government and government evaluation systems. Another key opportunity for international development partners to strengthen national evaluation systems is to support country-led evaluations as needed.

National Evaluation Capacity Index interpretation workshop with stakeholders in Guatemala

©María Fernanda López Vázquez

As in any other area of development support, coordination between development partners will contribute to the consistency of messaging, the avoidance of duplication or inconsistencies, and fostering of synergies for better results. Coordination with entities supporting other aspects of evidence-informed decision-making, such as the United Nations Department of Economic and Social Affairs or the United Nations Statistics Division, is also important. An assumption underlying the theory of change for evaluation is that reliable sources of data, including statistics and monitoring information, are available for evaluators to use.

It is also critical that development partners 'walk the talk', by supporting the development of an evaluation culture, promoting the use of norms and standards, sharing their evaluations, and using (and demonstrating that they are using) evaluations to support their own advocacy and decision-making processes.

This theory of change does not purport to be comprehensive. It is intended as a simplified overview of the elements needed to promote the use of evidence-informed decision-making. ECD support in any given context will require a more detailed theory of change, based on a context-specific initial assessment, with detailed pathways and analysis of risks and assumptions, and corresponding mitigation actions.

3. United Nations support to NECD and the evolution of national evaluation systems

This section draws on the survey, interviews, the country case studies and additional literature review to assess the support of United Nations agencies to NECD and the evolution of national evaluation systems. It starts by providing an overview of the support of United Nations agencies to NECD. Subsequent sections highlight United Nations support to elements of the national evaluation systems in the case study countries. These sections are structured according to key elements in the UNEG NECD theory of change (the enabling environment, institutional capacity and individual capacity), and the intended short- and longer-term outcomes leading to evaluation use.

3.1 United Nations support to NECD activities

UNEG members have been supporting a wide range of NECD activities, with support to individual capacity development being the most common. Request for support differs across countries depending on the level of development of the national evaluation system, ranging from support for training on evaluation basics to creating opportunities for putting evaluation policies into practice.

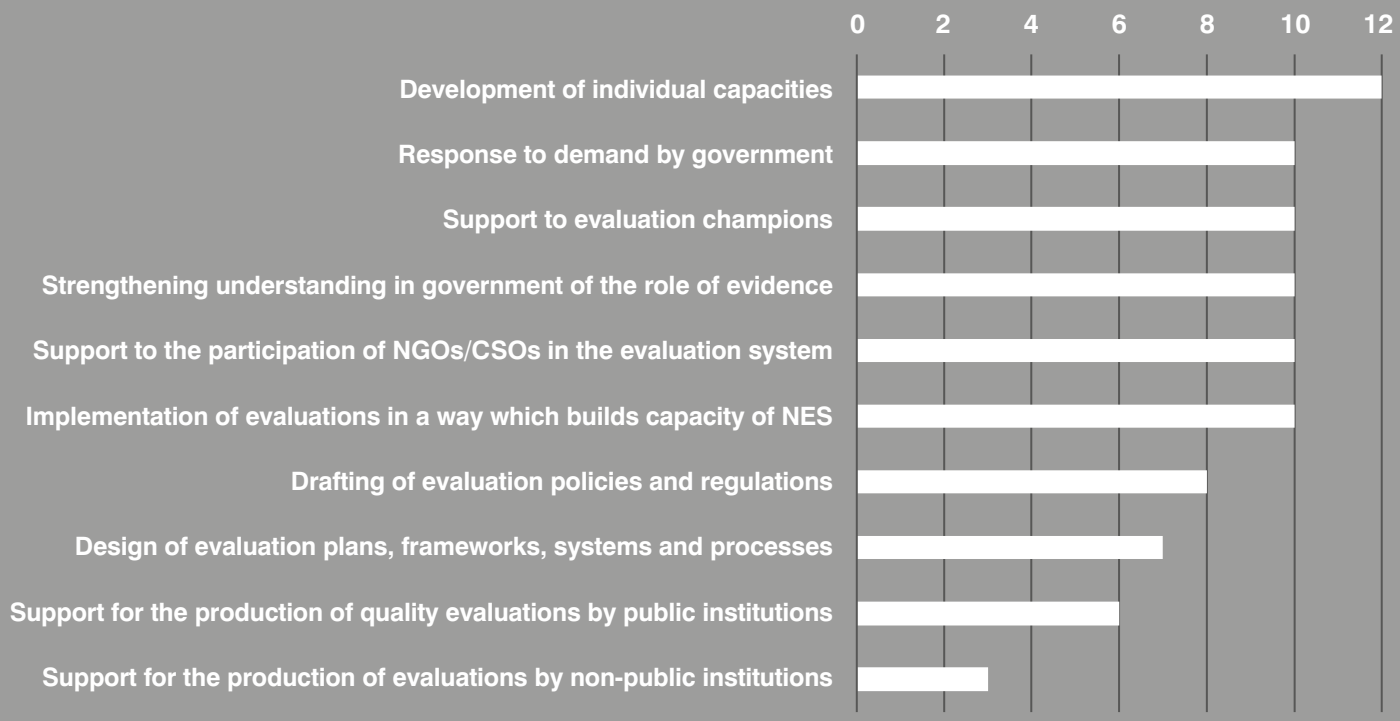
The survey of UNEG members was designed to identify and explore the type of NECD support United Nations agencies have been providing, specifically in the following areas:

1. Response to **demand for support** by government.
2. Supporting **evaluation champions** in the centre of government and more widely.
3. Strengthening **understanding in government of the role of evaluative evidence** in policy and practice.

4. Support for **drafting of evaluation or M&E policies and regulations**.
5. Design of **evaluation plans, frameworks, systems and processes**.
6. Development of **individual capacity to manage, conduct and use evaluative evidence**.
7. Support to the **participation of non-government stakeholders** in the national evaluation ecosystem.
8. Support to the **production of quality evaluations by public institutions**.
9. Implementation of **evaluations in a way which builds the capacity of national systems**.
10. Support to the production of **quality evaluations by non-public institutions**.

Fourteen agencies responded, two of which do not actively support NECD interventions. **Figure 2** provides the overview of responses to the survey (see Appendix 2 for the complete survey).

Figure 2. NECD support provided by United Nations agencies



Source: Survey of UNEG members

As Figure 2 illustrates, the largest contribution is to support the development of individual capacity, which was being done by most of those responding (12/14). Most responded to demand from governments (10/14), building champions in government (10/14), promoting the participation of non-government stakeholders in the system (10/14) and using their evaluations to develop capacity (10/14). A smaller group were helping to support the development of evaluation policies or regulations (8/14), frameworks or evaluation systems (7/14). Others (6/14) supported government evaluations, while only a few (3/14) supported evaluations produced by non-governmental organizations (NGOs). It should be noted that this does not necessarily capture all support to NECD as not all agencies responded, and responses may not capture all that the wider organizations (beyond the evaluation units) are doing.

Figure 3 shows which agencies were carrying out these interventions. UNICEF and UNFPA reported using all ten of these mechanisms, UNDP, WFP and UN-Habitat nine, UN Women eight, FAO and ILO six. UNITAR, as a training organization, only indicated developing capacity. OIOS which has an entirely internal focus, reported zero, with others reporting the use of a few.

Figure 3. United Nations agencies providing NECD support

UN Agency	Respond to demand from gov't	Build champions	Role of evidence	Policies/ regulations	Frameworks, systems	Individual capacities	Participation of non-gov stakeholders	Evaluations produced by gov't	Use evauations to build national systems	Evaluations produced by NGO	Total number of NECD areas
FAO	1	1	1			1	1		1		6
GEF	1	1	1	1	1	1	1		1	1	9
ILO	1	1	1			1	1		1		6
IOM				1			1				2
OIOS											
UN Habitat	1	1	1	1	1	1	1	1	1		9
UN Women	1	1	1		1	1	1	1	1		8
UNDP	1	1	1	1	1	1	1	1	1		9
UNFPA	1	1	1	1	1	1	1	1	1	1	10
UNICEF	1	1	1	1	1	1	1	1	1	1	10
UNIDO		1	1	1		1					4
UNITAR						1					1
UNODC	1					1			1		3
WFP	1	1	1	1	1	1	1	1	1		9
Total	10	10	10	8	7	12	10	6	10	3	

Source: Survey of UNEG members

National evaluation systems are at different levels of development, and therefore country requests for support vary. UNDP reported that country-level demand is greater for monitoring than evaluation support, and is often sectoral, rather than reflecting a "whole of government" demand. The East Asia and Pacific Regional Office (EAPRO) of UNICEF indicated that, in some countries with a level of established systems[17] and

17 Such as legal/ policy measures, budgeting and political commitments, with increasingly strong capacity in government departments and evaluation provider organizations.

evaluative culture (such as Malaysia or the Philippines), the emphasis is on putting policies into practice, seeking opportunities such as country-led evaluation processes. Other countries (such as Cambodia, Indonesia, Mongolia, Thailand and Viet Nam) demand more basic support, such as training, or developing M&E strategies. In countries where there has not been much previous engagement there may be an emerging appetite (such as China, Timor-Leste and the Pacific islands and territories). Demand may also depend on whether there is an expectation that support is available.

Box 5. Some examples of requests for support

ILO – M&E awareness

The national training initiatives conducted by ILO responded to the request from constituents (government representatives, workers' and employers' organizations) to know more about the use of M&E at national level to report on progress and key challenges around decent work and the SDGs. Special attention to the role of M&E activities was showcased, with examples of good practices within and outside the national territory.

UN Women – specific technical inputs

Governments were very interested in guidance and capacity development with regard to gender-responsive evaluation. There was also demand to receive technical assistance to support the evaluation of national strategies related to gender equality, or of thematic strategies on eliminating violence against women.

WFP – request for accompaniment

One WFP staff member sits in India's national Development Monitoring and Evaluation Office (DMEO) and has been working with DMEO staff on: i) National Evaluation Policy; ii) evaluation guidance and tools; iii) evaluation training; iv) learning seminars; and v) joint studies and evaluations.

IOM - interest in evaluations

IOM reported increased interest in evaluations related to migration, for instance within the framework of the Global Compact for Migration, or as a cross-cutting theme of the SDGs.

Source: UNEG NECD survey.

Some agencies are taking a more supply-driven approach, especially where their partner ministries are not active in evaluation. UNIDO reported that, so far, their NECD events have always been driven by them. For example, they proactively invited national partners to nominate participants for a workshop fully designed in advance by UNIDO. Ideally, this would be more demand-driven, with greater participation in the design of the workshops. Part of this can be explained by the still relatively limited role of evaluation in Ministries of Industry, their main counterparts.

3.2 Overview of the case study countries

Six country case studies were conducted to illustrate and strengthen understanding of the dynamics involved in developing a national evaluation system, and the roles played by United Nations agencies in supporting these systems. **Table 2** provides a brief summary of the status of each country's national evaluation system to facilitate reading of the remainder of the section.

Table 2. Outline of the state of the national evaluation systems in each country case study

Country	Status of the national evaluation systems
Benin	The national evaluation system was established in 2007, with a lead agency (BEPPAAG), national policy, guidelines, repository and 21 evaluations conducted through the national evaluation system. A law is being prepared. Postgraduate courses in M&E/evaluation are in place. There is one active VOPE.
Costa Rica	The system was legislated in 1994, but has been very active since 2014, with a national champion in the Ministry of National Planning and Economic Policy (Mideplan), two evaluation plans with over 27 evaluations undertaken or planned, and a repository of these evaluations. A policy is in place and an active consultative body with stakeholders on evaluation. Postgraduate courses in evaluation have been in place for some time.
Kenya	The M&E process has been active since the early 2000s, and a policy was approved in 2012, with a new version approved by Cabinet in May 2021. Evaluation guidelines have been produced, but no evaluations are being undertaken by Government at present, except for midterm reviews of devolved county development plans. There is one active VOPE and many postgraduate courses in M&E/evaluation.
Morocco	Evaluation is enshrined in the Constitution and decrees, but there is no policy. There is a partial national evaluation system, with the National Observatory of Human Development (ONDH) tasked with conducting evaluations but not across the whole government, and multiple evaluations have been conducted. Postgraduate courses in M&E/evaluation are in place. There is one active VOPE.
Philippines	A policy was approved in 2014, and the National Economic and Development Authority (NEDA) is the lead agency in Government. Some evaluations have been undertaken, although there is no national evaluation plan or agenda. Some guidelines have been developed, but not yet formally issued. There are no postgraduate courses in M&E/evaluation. There are two active VOPEs which bring together consultants and academics, and a government M&E learning network. Two senators are introducing different M&E bills in the Senate.
Sri Lanka	Sri Lanka has an approved policy and a champion (Department of Project Management and Monitoring, DPMM). There was a functioning system in the early 2000s, but no evaluations are being carried out by Government at present. A Policy Framework to implement the policy is with Cabinet for approval. There is a draft Bill in Parliament, which historically has been a strong advocate of evaluation. There is one very active VOPE, and the only postgraduate course in M&E in the region.

3.3 Fostering an enabling environment

As mentioned in section 2, an enabling environment for evaluation includes, on the one hand, the political will or motivation to produce and use evaluations and, on the other, the broader governance environment, including the policy and regulatory environment as well as engagement between public and private stakeholders. This section examines these different aspects with reference to the case study countries.

3.3.1 Policy and regulatory environment

The focus on results-based management (RBM) introduced by development partners while tracking the performance of their own programmes helped to raise government interest in evaluation. The development of evaluation policies and legislation has been supported by United Nations agencies in all case study countries except Costa Rica. Even where policies are in place, implementation is still a challenge.

A central element of the NECD theory of change is that a favourable policy and regulatory environment will support the use of evaluation. In many countries, the move towards RBM and performance budgeting approaches has helped to promote a culture where evaluation is recognized and valued. This has also been driven through the efforts of development partners to track the performance of their own projects and programmes, while at the same time building M&E capacity in their projects and the public sector (e.g. Trikawalagoda, 2018). This is apparent from most of the case studies.

Round table "Institutionalizing evaluation in the agriculture sector" organized by FAO and EvalForward at the Francophone International Forum on Evaluation - FIFE2019, Burkina Faso

Table 3 shows that all case study countries had national evaluation policies in place, except for Morocco. In Costa Rica, the policy was only approved in 2018, long after the system had started operating. Only Costa Rica had a law underlying the M&E system, while Morocco had a clause in the Constitution and some decrees. The Philippines and Sri Lanka have draft bills which may be passed, and a law is being drafted in Benin. In most cases, countries benefited from the support of United Nations agencies, in particular UNICEF and UNDP, as well as other donors.

Table 3. Status of evaluation policy or legislation in the case study countries

CASE ELEMENT	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica
M&E/Evaluation Policy adopted	2015 (NEPF)	2019 (NEP)	2012	2012		2018
M&E Act	Draft Bill, likely to be passed	Draft Bill, may not be submitted		Act being prepared	Constitution plus other decrees	Range of legislation, not specifically M&E
Role of United Nations agencies	UNICEF with NEPF, now Bills	UNICEF with NEP, Bill	UNICEF on policy	UNDP on policy	UNDP on policy/ legislation	
Other				World Bank supporting new Law		

Key: NEPF = National Evaluation Policy Framework, NEP=National Evaluation Policy.

- In the **Philippines**, public sector evaluations are governed by the National Evaluation Policy Framework (NEPF), jointly issued by NEDA and the Department of Budget and Management in 2015. UNICEF supported the two agencies to develop the Framework. As approved, the NEPF applies across government, but does not have the status of an Act. At the time of researching this report, UNICEF was supporting parliamentarians, and two separate bills were being tabled in the Senate.
- The planning system in **Costa Rica** was initially grounded in the National Planning Law of 1974, which includes a mandate to systematically evaluate programmes, plans and policies. However, the national evaluation system was not initiated until 1994, with a law conferring the responsibility to coordinate, evaluate and monitor programmes and development policies to Mideplan. Only in 2018 was a National Evaluation Policy (NEP) formulated to guide government evaluation, when the national evaluation system had already been under implementation for some time.

The United Nations did not play any specific role, as the process was mainly pushed at national level.

- A range of development partners have supported NECD in **Sri Lanka**, starting with UNDP and the Asian Development Bank (ADB) in the late 1990s. UNICEF has also been supporting NECD in Sri Lanka for several years, in particular for the development of the NEP. In 2016-18, UNICEF supported consultations and advocacy for the policy, and the establishment of a Parliamentary Forum where the two motions to the Parliament were submitted: first, to establish an NEP; and second, to allocate public funds for evaluation.
- **Kenya** launched its National Integrated Monitoring and Evaluation System (NIMES) in 2004 with the establishment of the Monitoring and Evaluation Directorate (MED) in the Ministry of Planning and National Development. NIMES is used to track implementation progress of the Kenya Vision 2030 and, since 2010, of the County Integrated Development Plans. MED has organized annual M&E weeks since 2012 and this helped to raise awareness of M&E in the public sector, including the potential of evaluation. A national M&E Policy was approved in 2012, and a revised version was tabled in 2017 and is awaiting approval. Overall, a broad ecosystem for evaluation has developed, but evaluation is not undertaken systematically in government. United Nations agencies (UNDP and UNICEF) played an important role in the development of the national evaluation systems alongside others such as the World Bank, United Kingdom Department for International Development (DFID), United States Agency for International Development (USAID) and the Swedish International Development Cooperation Agency (SIDA).
- In **Benin**, the process of institutionalizing evaluation started gradually, with the election of a new Government in 2006, with a vision to make evaluation a major governance tool, and which set up BBEPAAG in 2007. The NEP was established in 2012 and defines the overall framework for planning and carrying out evaluations. The NEP was evaluated in 2019 and this led to an action plan to adjust implementation over the period 2020-2021. Progress in the national evaluation system has mainly been promoted by development partners, notably UNDP and UNICEF.
- In **Morocco**, His Majesty King Mohammed VI launched the National Initiative for Human Development (INDH) in 2005, to improve social welfare. This was accompanied by a renewed interest in accountability and M&E, highlighted in successive Royal Speeches from 2001, and led to a decision in 2009 to mandate the ONDH to evaluate the INDH. Another milestone was reached in 2011 with the promulgation of a new Constitution, which provides for public policy evaluation (PPE). In parallel, the Government drafted the Advanced Regionalization Report of 2011 intended to reduce territorial disparities and enhance competitiveness, among others. The subsequent Organic Laws of 2015 mandated elected Regional Councils to report on the outcomes and impacts of plans, programmes and projects to the Regional Comptroller Courts. However, there is no single NEP. United Nations agencies support ONDH, the High Commission for Planning, INDH, the Economic, Social and Environmental Council, and other line ministries.

Of the 14 United Nations agencies who responded to the survey, eight indicated that they had supported policies and regulations. Beyond the case study countries, UNDP reported that they had supported work on evaluation and M&E policies in Botswana, Malawi, Madagascar and Zimbabwe, and a law in Bosnia and Herzegovina, often in collaboration with UNICEF. UNICEF reported that they had worked on policies in Nepal, Pakistan, the Philippines, Thailand and Zimbabwe, among others. WFP has seconded an M&E expert to the Indian Government to help establish an M&E policy and strategy, among other elements. In 2020, UNFPA supported the Asia-Pacific regional dialogue on national evaluation policies and systems, where participants learned from other countries and received technical knowledge on NEPs.

3.3.2 Role and engagement of stakeholders

Non-governmental stakeholders, including VOPEs, academia and NGOs, play a critical role in making an evaluation ecosystem function effectively. In most case study countries, VOPEs contributed by raising awareness on evaluation, helping to move the evaluation agenda forward. Only Costa Rica has a formal national evaluation platform that facilitates inclusive dialogue around national evaluation issues, while in the other countries the participation of non-government stakeholders is on a more *ad hoc* basis. The case of Benin shows how inclusive evaluation processes can result in significant institutional reforms.

The NECD theory of change considers "stakeholders exchanging knowledge and coordinating around the national evaluation agenda" as part of an overall enabling environment that would help ensure that evaluations are produced and used. This includes supporting CSOs and research institutes to participate in steering mechanisms for national evaluation systems.

Within United Nations evaluations, non-government stakeholders play a role in steering committees, reference groups, validation workshops and dialogue processes related to evaluations in most of the countries.

Only one case study country has a formal mechanism included in the policy for stakeholder dialogue, the National Evaluation Platform of **Costa Rica**. The Platform was initiated through the 'Evaluation Capacity Development in Latin America' (FOCEVAL) project, implemented by Mideplan and DEval. Through this Platform, representatives of the National Assembly, the Ministry of Finance, the Auditor General, academia, civil society and other bodies meet twice a year to discuss national evaluation issues, with additional sessions as needed. The meeting agenda covers: accountability for political goals as per National Development Plan priorities; exchanges of experiences and evaluations implemented by stakeholders; and specific issues proposed by participants. Informants considered this Platform a good practice to guide the national evaluation process. In all other countries, the participation of non-government stakeholders is more *ad hoc.*

The **Benin** case study shows CSOs and NGOs very active in evaluation processes, participating constructively in evaluation steering committees and playing a significant advocacy role. In 2009, BBEPAAG carried out an evaluation of the agricultural sector policy. The steering committee was made up of key stakeholders including CSOs, and chaired by the Ministry of Planning. The report was validated through a three-day workshop bringing together the main stakeholders of the sector, who were also involved in revising the policy. This inclusive process allowed the evaluation to initiate significant institutional reform.

VOPEs can have a significant influence on the wider evaluation agenda. In the survey, UNDP, UNICEF and WFP indicated that they had supported VOPEs, and this support was significant in the cases of Sri Lanka and Kenya. VOPEs can play a role in raising awareness on evaluation and advocating for the establishment of national evaluation systems. For example, the Global Parliamentary Forum for Evaluation, held in Sri Lanka in September 2018 and supported by UNICEF, enjoyed the strong involvement of VOPEs, and was influential in advocacy around evaluation, notably for Sri Lanka and the Philippines. In 2019, the Asia-Pacific Evaluation Association Conference held in Manila was organized by a VOPE from the Philippines, and was similarly influential with the Philippines Senate and Congress.

In Benin, Kenya, Morocco and Sri Lanka, VOPEs are involved in annual or biannual evaluation weeks, and have often been the torchbearers for evaluation through periods when government interest waned. For example, the Moroccan Evaluation Association was created in 2005, and since then has been advocating for the institutionalization of evaluation in the country. A series of memoranda have been presented to institutions with recommendations on how to integrate PPE into the wider system through a sound regulatory framework, decentralization and developmental plans.

In **Morocco**, triangular relationships between academia, the lead government evaluation champion and development partners have been important. Partnerships between the Moulay Ismail University Masters in Evaluation, the United Nations and ONDH provide an example where each institution reinforces the work of the others, helping to lay the groundwork for further development of the national evaluation system.

3.3.3 Evaluation champions and persistent political will

Developing a system is a long-term process which requires persistent political will and strong advocacy. Evaluation champions are instrumental in advancing the national evaluation agenda. Champions can be located in parliament or the executive, and can include VOPEs or academia. Having more than one champion and building a coalition is important to support change. Progress is vulnerable to political transition and thus having legislation in place can create more stability.

The enabling environment includes political will, which is often a challenge, and not always easy to influence by projects or programmes designed to strengthen evaluation capacity. In countries like Benin and Costa Rica where the national evaluation systems

have advanced rapidly (as well as Uganda, South Africa and Mexico that were not covered by this study), there has been strong political will and advocacy. In Benin, the national evaluation system is the result of the political will of the Government elected in 2006, which established BBEPAAG with the aim of having effective public policies capable of improving the living conditions of citizens. In Costa Rica, the national evaluation system began to develop in 1994, getting gradually stronger through institutional agreements to monitor and evaluate strategic projects, reinforced in a reform of the Constitution in 2000, and from 2005 evaluation started to be taken more seriously. Another incentive was added with the country's announcement in 2012 that it intended to join the Organization for Economic Co-operation and Development (OECD).

Evaluation champions are key to creating interest in evaluation and the development of evaluation systems. Champions might be located in parliaments, M&E units, VOPEs, CSOs or academia. In all the case study countries, UNDP and UNICEF are the main United Nations agencies working with institutional champions, both in the executive and parliament. In the survey, most agencies reporting support to evaluation champions (10/14) did so by identifying individuals and building their profile (4/14), exposing them to good practice in the country (4/14), or outside the country through reports and other means (6/14), or through visits (3/14), or by providing training (9/14).

Table 4. Evaluation champions

	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica
Champion in Parliament	Senate and Congress	Was strong, not currently	Parliamentary caucus	No	Yes	No
Other champions	NEDA	Strong VOPE, very involved with EvalPartners	Active VOPE, some active county governments		Moroccan Auditor General, High Commissioner for Planning	Strong champion in University of Costa Rica
United Nations support to champions	UNDP (NEDA), UNICEF (Senate/ Congress)	UNDP, UNICEF	UNICEF supported MED and VOPE	UNDP, later UNICEF	UNDP, UNICEF, UNODC, UN Women	UNICEF active in Platform.
Other support to champions	3ie/DFAT		SIDA, CLEAR-AA, Twende Mbele	Twende Mbele	Westminster Foundation for Democracy	Germany supporting Mideplan and University of Costa Rica

Facilitating emerging evaluation champions in parliaments can contribute to the push for an evaluation policy or law. In Sri Lanka, following the participation of parliamentarians in regional events and the establishment of the Parliament Forum supported by UNICEF, motions were put forward for the NEP with funding allocations in the national budget. In the Philippines, UNICEF work with Congress and the Senate has resulted in two bills on M&E being tabled in the Senate.

The case studies also illustrate potential vulnerability to political or administrative transitions. In Sri Lanka, for instance, strong evaluation champions emerged in Parliament between 2015 and 2019. A draft bill was tabled in Parliament in 2019 to enact the NEP, but has been delayed by changes in Government. Passing a law is likely to create more stability across political or administrative transitions, but as the example of Mexico shows (not a case study country), even where an Act was in place, the evaluation agency (CONEVAL) was still vulnerable to changes and in 2019 its budget was cut drastically.

To make the system less vulnerable to political and administrative transitions, it is important to build a coalition to support change, and support more than one champion over time. In Benin, the case study identified several champions over the years, including the former President who in 2007 created a specific Ministry for Evaluation, and in technical ministries (such as the former Deputy Minister of agriculture). With this long-standing political commitment and the established system, despite some shifts over time, the Government is now funding a substantial part of evaluation costs. In Costa Rica, the United Nations and others have supported the role of Mideplan as champion of the national evaluation system and the institutional stability of the country has contributed to the development of the practice and institutionalization of evaluation.

3.4 Institutional capacity

3.4.1. Evaluation institutions established

The presence of a centrally-situated evaluation function (e.g. in a ministry responsible for planning) can facilitate the development of a national evaluation system. Ideally this should be an agency with some authority over line ministries to better play an oversight role, but they need to work in such a way to build buy-in across government. It is also possible to start a system by working with a technical line ministry if there is interest in engaging in the process.

Key to creating interest in evaluation and the development of evaluation systems is the presence of an evaluation (or M&E) function or unit located either centrally, for example in a ministry responsible for planning, or in line ministries. In most cases, the main institution in charge of evaluation is in the executive, often situated in the treasury (e.g. Chile, Kenya) or planning department (e.g. Costa Rica, Lesotho), with the presidency/office of the prime minister (Benin, South Africa, Uganda), or at times as an independent institution (e.g. Mexico, Morocco). However, in practice, in many countries

M&E remains fragmented. For example, in the Philippines, there is a proliferation of M&E functions within agencies, with different offices in the same department undertaking M&E of their programmes.

Table 5. Main institutions in charge of evaluation

	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica
Main institutions in charge of evaluation	National Economic Development Authority (formerly also DBM)	Department of Project Management and Monitoring in Treasury and Planning	M&E Directorate, in Treasury and Planning	BEPPAAG, initiated in 2007 under the General Secretariat of the Presidency of the Republic	National Observatory of Human Development, line ministries, agencies	Ministry of Planning and Economic Policy

A central institution in charge of evaluation is key, ideally where the agency has a position with some authority over line ministries, enabling it to play an oversight role and coordinate evaluation efforts across different sectors. For instance, the Sri Lanka case study notes that DPMM, the government institution in charge of evaluation, has never been in a very strong position, with inadequate capacity and resources to effectively promote NECD. Some suggest that DPMM would be better positioned above line ministry level, allowing it to more easily carry out its oversight function.

3.4.2 Institutional mechanisms for evaluation in place

Translating evaluation laws and policies into practice requires guidelines, standards, evaluation plans, etc. While these multiple elements of the system need to be built progressively, it is possible to start to conduct evaluations before they are all in place. The availability of human and financial resources is also essential. The case studies show that funding from development partners can be catalytic and leverage additional resources from government.

While a policy can provide an overall framework for evaluation, in practice it is the detailed guidelines and systems that translate policy into practice. All of the case study countries have developed evaluation guidelines, with support from UNICEF (Benin and Kenya), UNDP (the Philippines) and Westminster Foundation for Democracy (Morocco). Evaluation standards only exist in some countries, for example in Kenya with support from UNFPA. In Sri Lanka, evaluation standards and guidelines were developed by the Parliamentarians Forum for Development

ILO training activities, China

Evaluation through a series of consultations and workshops, and these are expected to be reviewed and adopted by DPPM, which is charged with preparing an NEPF and implementing mechanisms.

A funded national evaluation agenda or plan is an important mechanism to materialize government demand for evaluation. Costa Rica is the only case study country that has developed a formal national evaluation agenda. In the Philippines, UNDP is supporting the development of a toolkit for drafting a National Evaluation Agenda.

Another element of the system is the evaluation competency framework, which contributes to the professionalization of evaluation. In the Philippines, UNDP has supported NEDA to develop an evaluation competency framework and capacity development plan. UNFPA has supported the development of a competency framework for evaluators in the Asia-Pacific region, especially targeting young and emerging evaluators, as well as a mentoring programme.

The NECD theory of change postulates that making evaluations publicly available is important in fostering their use. Benin and Costa Rica have a repository of evaluations, which are accessible online, while Sri Lanka's web-based repository is no longer functioning. Searches of key websites in the other four countries suggested that, even where evaluations are led by donors with government partners, they are not seen as government resources to be made publicly available.[18] Of the six United Nations agencies supporting government evaluations, four indicated that these evaluations were made public (for one as part of the funding agreement), for three on government websites and for four on United Nations websites.

18 To test whether evaluations were available online on ministry websites in the other four countries, they were searched on the websites of the Departments of Agriculture, Health, and Social Development. None were available on these websites, even though some other research studies were found.

Table 6. Presence and support for evaluation systems in case study countries

CASE / ELEMENT	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica
Evaluation plan	Planned		Started to develop		ONDH, not national	National Evaluation Agenda
Guidelines	Not formally issued	Just developed	Kenya Evaluation Guidelines 2020	National Methodological Evaluation Guidelines	First PPE guide	Guidelines
Standards		Basic	Supported by UNFPA		ONDH not national	
Competency framework	Underway					In place
Evaluation repository	Designed	Existed in the past		In place		In place

In the survey, seven United Nations agencies indicated that they had supported frameworks and systems, all of whom had supported the development of evaluation guidelines, four the development of evaluation plans or agendas, and five the elaboration of evaluation standards and competency frameworks. For example, UNDP Nepal supported the development of National Monitoring and Evaluation Guidelines through the 'Strengthening National Planning and Monitoring Capacity' project (UNDP IEO, 2021). UNDP Mexico is working on the adaptation of evaluation methodologies to incorporate Agenda 2030 principles for the evaluation of programme processes, results and impacts. They are also supporting the Technical Secretariat for Planning and Evaluation of the Government of Yucatán to: strengthen design evaluation exercises based on the results of a meta-evaluation; implement a strategy to link the results of M&E processes with budget decisions; and, incorporate the Agenda 2030 approach into their Evaluability Strategy (UNDP IEO, 2021). In Niger, UNDP supported the creation and operation of a M&E Unit for the national Economic and Social Development Plan (UNDP IEO, 2021). UNICEF has supported Kenya, Pakistan and Zimbabwe to develop guidelines, and WFP is providing technical support to the Tunisian Government to draft an M&E framework. The M&E expert seconded by WFP to DMEO in India also works on guidelines.

Box 6 describes an ambitious joint programme for M&E support to the Government of Malawi, an older example which was not without challenges, but which illustrates the type of joint programme which is possible in the M&E space.

Box 6. Joint programme supporting M&E in Malawi

The United Nations in Malawi provided support to the Government through the Joint Programme Support for Strengthening Monitoring and Evaluation Systems in Malawi (2008–2013), financed through a basket fund with contributions from the European Union, UNFPA, UNDP, UNICEF, DFID, GIZ and the Government of Malawi. The objective of the programme was to strengthen and develop sustainable national systems for M&E of development strategies and programmes. The programme adopted a whole-of-government approach to capacity development, with support to the policy and institutional framework, a master plan for M&E, and the establishment of M&E positions in all 28 district councils of the country as well as M&E frameworks for 23 sectoral ministries. Training was provided in M&E, project evaluations, impact assessment and policy analysis. Support was provided to constitute a national VOPE, develop a university training module on RBM, and create tools such as an RBM handbook by the Ministry of Economic Planning and Development and an important study on the state of M&E in Malawi in 2014, which became a reference document for articulating M&E priorities in the public sector. UNDP has also supported the development of a draft Monitoring and Evaluation Policy.

Source: UNEG NECD survey.

Another element related to institutional capacity is human and financial resources. Historically, middle-income countries like Costa Rica, Morocco, the Philippines and South Africa have been able to pay for government evaluations, although the fiscal situation during the COVID-19 pandemic has made this much more difficult. In Morocco, ONDH contributes significantly through cost-sharing agreements for joint programmes supported by the United Nations.[19] Where political commitment exists and the system is established, even in low-income countries like Benin (though classified since 2020 as lower middle-income), governments are funding a substantial part of evaluation costs. Interviewees indicated that around 60 percent of the cost of government-led evaluations is now funded by Government (although relatively few evaluations are conducted), where previously their evaluations were fully funded by donors. In Sri Lanka, the proportion of financial resources dedicated to evaluation as compared to implementation is still very marginal, there are no specific budget lines for evaluation, and allocations to ministries to support their M&E units are limited.

In summary, many projects are working to build the institutional elements of their national evaluation systems, but of the case study countries only Costa Rica, and to some extent Benin, has many of these elements in place. In the other countries, the national evaluation systems have elements in place but do not yet constitute working systems across government. These multiple elements of the system need to be built progressively, and do not need to all be in place to start conducting evaluations, as discussed in the next section.

19 US$ 1 million of $1.5 million for HEPP Phase 2, and $1.75 million of $4 million for the ONDH Support Programme.

© Asela Kalungampitiya

EvalColombo 2018 conference

3.4.3 Strengthening national evaluation systems through United Nations evaluations

The limited number of evaluations conducted with active involvement of government results in limited opportunities for government officials to practice and develop their capacity. United Nations evaluations can be conducted with governments to pilot how their evaluation systems can be developed. As such, conducting joint and country-led evaluations using national systems can contribute to enhancing institutional capacity, while increasing ownership and the use of findings.

There is potential to considerably strengthen the impact of evaluations conducted by United Nations agencies, especially where evaluation systems exist, by using government terminology and systems to implement the evaluations. However, only two agencies (UNICEF and WFP) reported sometimes using government terminology and systems to implement their evaluations. For example, UNICEF cited using government evaluation terminology and systems to undertake the evaluation of Zambia's 7th National Development Plan. Only these two agencies reported using government procurement systems, although WFP specified that this was rare. In Costa Rica, Mideplan reports that 33 percent of international organizations use their systems.

UNICEF, WFP and UNDP indicated that they have used evaluations with governments to pilot how their evaluation systems could be developed. The UNICEF-funded evaluation in South Africa described in **Box 7** provides an example of this, and shows how it can have significant impacts. This example was shared in many interviews, and suggested as a possible viable way forward with widespread acknowledgement that it would be a very positive approach to pilot in other countries.

Box 7. Using evaluations to pilot national evaluation systems: collaborating with UNICEF in South Africa

In July 2011, South Africa undertook a powerful study visit to Colombia, Mexico and the United States of America with the Deputy Minister and Director-General of the Department of Performance (later Planning), Monitoring and Evaluation (DPME) and other departments to look at their evaluation systems. During the study tour an outline for conceptualization of the South African system was agreed, and it was decided to pilot an evaluation, as learning-by-doing, to work out how the national evaluation system should be set up. On return, in August 2011, the team developed the NEP in a write-shop, which went for public consultation and was approved by Cabinet in November 2011. During the consultation, the topic of Early Childhood Development was agreed for a pilot evaluation, terms of reference were developed, and the pilot started in October 2011. This was partly funded by UNICEF, who sat on the steering committee, and partly by the Government who also managed the procurement. The report was approved in June 2012. This evaluation was used to pilot how the evaluation system would run, and the guidelines were the starting point for developing aspects such as the terms of reference, the inception phase, management response and the improvement plan. This was a classic example of how just implementing an evaluation can help build the government evaluation system. Further information can be found in Davids et al. (2015).

Source: Ian Goldman.

In most of the countries reviewed, the low number of evaluations conducted with active involvement of government officials has limited opportunities to strengthen institutional capacity to conduct, commission and use evaluations. For example in Sri Lanka, due to the small number of practical evaluations involving government officials, there have not been many opportunities for learning how to plan and commission evaluations. There is demand for hands-on evaluation training, through joint evaluations and on-the-job learning. In this context, it is interesting that only six United Nations agencies indicated having supported government evaluations with funding, technical assistance or participating in steering committees. UNDP, UNFPA, UNICEF, UN Women and WFP all provided examples of evaluations they had supported. For example, in Burundi UNICEF supported the development and conduct of a country-led evaluation in 2020–21 (on basic education reform) and is now supporting the development of a country-led evaluation on malnutrition for 2022. It is also supporting an SDG evaluation on Social Protection in Madagascar.

One respondent noted that, ideally, all United Nations evaluations should be country-led or at least conducted jointly. However, in practical terms, the ability to do so may depend on agency in-country presence to facilitate this engagement, the commitment of agency management, and the country context. Another respondent echoed this, noting that in-country capacity was necessary in order to support country-led

evaluations, while acknowledging that this should be a priority "16 years on from the Paris Declaration." The same respondent queried: "how much is the international development system ready to change?" Box 8 describes an example of a country-led evaluation of United Nations support in South Africa in 2009, the first of its kind.

Box 8. Joint evaluation of the role and contribution of the United Nations system to the Republic of South Africa

Conducted between September 2008 and March 2009, the evaluation assessed the relevance and effectiveness of cooperation between South Africa and the United Nations system within the three-tier strategic policy priorities of the country: a better South Africa, a better Africa, and a better world. The evaluation was unique for a number of reasons:

- **The Government of South Africa expressed the will to develop a policy dialogue to strengthen its partnership with the United Nations based on evaluative evidence.**
- **This was the first time that the United Nations system as a whole had been jointly evaluated at the country level, rather than on an agency-by-agency basis.**
- **Building trust and sharing the will to improve, based on lessons from past experience, were essential aspects of the exercise. All important decisions were made by consensus.**
- **The evaluation demonstrated the need for champions, and there was clear leadership on both sides of the partnership.**
- **Key to the success of the joint evaluation was the fact that it was conducted by a highly competent and independent evaluation team that had no conflict of interest with the United Nations system or the South African Government. The Joint Evaluation Management Group, comprising evaluation specialists from South Africa and UNEG was also independent from line management functions on either side.**

Source: http://www.unevaluation.org/document/detail/284

While it is not always possible to realize country-led evaluations, it is essential to meaningfully involve government from the outset of an evaluation to enhance ownership and use of results. This was stressed, for example, by government officials in Sri Lanka who listed resistance to donor-driven or independent evaluations commissioned by the donors among the main factors negatively affecting the use of evaluations in the public sector.

Several agencies responding to the survey indicated the involvement of national actors in reference groups, workshops, etc. One respondent reported: "When we started country programme evaluations … we started to engage governments in these

evaluations, so United Nations agencies serve government purposes better. We often found some policymaker who showed interest. We involved staff in design workshops, briefings of the evaluation team, co-organizing workshops – expanding outreach beyond beneficiaries for example to religious communities and political parties, to have a more constructive approach."

3.4.4 Institutional capacity: non-public sector organizations

The case study countries show different degrees of capacity and organization of non-public stakeholders, which influences their capacity to engage in national evaluation systems. Some United Nations agencies have directly supported VOPEs, and most agencies involve non-state actors in their own evaluations and capacity development activities. Morocco also shows how CSOs can be a driving force for evaluation at decentralized levels.

The NECD theory of change proposes that the supply and demand of evaluation capacity beyond government institutions (both in terms of evaluators and ECD offerings) is important for the national evaluation ecosystem. Non-governmental stakeholders have several roles to play in developing this ecosystem, from providing training (e.g. universities) and building the evaluation profession (VOPEs), as evaluators (universities, consultants, think tanks), as clients, peer reviewers of evaluations, stakeholders with different perspectives on the programmes or policies of focus, or as a source of demand for evaluation.

The different country case studies showed different levels of organization, capacity and engagement by non-public sector actors in the national evaluation systems. In Costa Rica, the National Evaluation Platform is a very important structure steering the system, and membership includes civil society and academia.[20] In Kenya, there are strong research think tanks like the Kenya Institute for Public Policy Analysis, a Parliamentary Caucus for Evidence, various universities and other institutions offering academic and professional trainings and research, and a recognized VOPE, the Evaluation Society of Kenya (ESK).

The Sri Lanka Evaluation Association (SLEvA) was established in 1999 and comprises experts from academia, international agencies, non-governmental agencies, public and private sectors. SLEvA has been a driving force for knowledge exchange, holding a range of events, conferences and training activities that have been instrumental in fostering an evaluation culture in Sri Lanka. In collaboration with UNICEF, the University of Sri Jayewardenepura established a Centre for Evaluation, the first professional evaluation institute in South Asia.

Generally, however, there was not much evidence for demand for evaluation from non-public sector organizations from the case studies or survey. Librado & Maclean (2019)

20 Uganda has a similar structure. South Africa instead has a strong VOPE which interacts a lot with government, but civil society does not have any formal steering role in the national evaluation system.

found an increase in demand for evaluation from citizens in seven Asian countries, as well as in mechanisms facilitating citizen engagement in evaluation processes, and concluded that progress on government openness is mixed, but advancing overall. In Morocco, local-level demand for evaluation has increased due to the advanced decentralization process, among other reasons. There are civil society demands to complete the policy cycle with evaluation, which would eventually lead to the improved effectiveness of programmes. For example, the Tensift Regional Development Centre is an organization supporting regional development policies and the evaluation of public policies in the Marrakech Safi region. The University of Costa Rica also reports that there is local-level demand for evaluation originating from municipalities, through master's students who use the cases for their final theses. In general terms, demand for evaluation from the University of Costa Rica has increased recently to one or two requests per week.

Ten United Nations agencies reported supporting the participation of NGOs in various mechanisms: structures overseeing evaluation in the country (3), steering committees (9), events guiding the development of evaluation terms of reference (7), validation workshops on the results of evaluations (9), and dialogue processes around evaluations (7). WFP, UNFPA, UNDP and UNICEF reported supporting VOPEs with capacity development activities. For example, the UNFPA EAPRO supported virtual awareness sessions for regional VOPEs on the professionalization of evaluation and use of the competency framework.

Ada Ocampo, Senior Evaluation Specialist, UNICEF Evaluation Office and Riccardo Polastro, Regional Evaluation Advisor, UNICEF East and the Pacific Regional Office. Manila, Philippines Senate

ILO works with stakeholders including workers and employers, and so involves them in training and in the evaluations themselves. WFP and DEval have collaborated on a National Evaluation Capacity Index analysis in Latin America (INCE) involving VOPEs and academia from the start. WFP has also been using a tool called EvaluVision to apply visual thinking in validation workshops and increase ownership of the evaluation process by non-governmental stakeholders, particularly at the community level in South-East Asia (WFP, 2021). UNICEF Eastern and Southern Africa Regional Office has supported the inclusion of NGOs in structures overseeing evaluation in Burundi and Madagascar, as well as dialogue processes on evaluations.

3.5 Individual capacity

This section uses two lenses to look at individual capacity: firstly, how individual capacity is being built and supported (e.g. through training and other programmes), and secondly, how awareness of the relevance of evaluation is being strengthened. It also presents South-South and triangular cooperation, including international networks, which can contribute to individual and other capacity.

3.5.1 Developing individual capacity to manage, conduct and use evaluations

All responding agencies supported training, but fewer supported other capacity development interventions such as learning-by-doing or mentoring. Other initiatives to strengthen individual capacity include postgraduate courses, which emerged as very important to develop a group of people with in-depth training on M&E or evaluation, and communities of practice. It is important to combine individual and institutional capacity development, as trained individuals will often move on.

The most commonly reported form of United Nations support to NECD is through training. All 12 agencies that reported supporting NECD indicated they used training, while two reported facilitating study visits, five learning-by-doing, five mentoring, six mentioned providing guidance and three were providing internships. Some were aiming to create synergies and embed M&E into existing capacity development programmes without a strict M&E focus.

In terms of actual training delivery, UNITAR (a specialized training organization) reported combining e-learning and face-to-face training with mentoring. UNICEF provided evaluation training to government officials in Rwanda in 2019. UNICEF supported the development and delivery of an M&E course (university level, including an impact evaluation module) through a training-of-trainers model in Zambia in 2021. UNDP country offices support various trainings on RBM, M&E, and sometimes specifically on evaluation. The International Fund for Agricultural Development (IFAD) collaborated with CLEAR centres to develop the 'Programme in Rural Monitoring and

Evaluation' training framework. This M&E curriculum tailored to the rural development sector is designed for national counterparts to use with internationally-financed projects to better respond to the needs of the sector. WFP provided training on evaluation to government staff from India, Namibia, Tunisia and Malawi, as well as specific training on impact evaluation in Kenya. FAO, WFP, UNODC and UNDP also mentioned supporting people to attend conferences, including the NEC Conference pre-event training workshops. Some agencies are now supporting young emerging evaluators, including UNFPA and UNICEF.

There are postgraduate courses in M&E or evaluation in all of the case study countries except for the Philippines, often supported by UNICEF (or UNODC in the case of Morocco). According to informants to the Morocco case study, United Nations engagement gave the master's course and the University an international dimension that did not previously exist, and has attracted more interest from students and professors. In Costa Rica, the active role played by a university with peer linkages with a European university has strengthened the evaluation system. They have also developed linkages with other universities in the region, notably Ecuador, to enhance South-South cooperation in the area of evaluation and disseminate the Costa Rican experience. The absence of postgraduate courses in the Philippines was considered quite problematic by some respondents.

In terms of learning networks, FAO, WFP and IFAD supported peer-to-peer exchange and learning through the EvalForward Community of Practice, with over 1,000 registered members at the time of writing, and which organizes regular online discussions, webinars, EvalForward Talks (where members share evaluation-related challenges or experiences) and other knowledge sharing activities. In the Philippines, a year-long webinar series on evaluation is taking place to augment capacity development activities, which is also expected to support efforts to build an active community of practice for evaluation in the country. Other initiatives with links to VOPEs were reported in the survey.

Several examples of learning-by-doing emerged. Of the 14 agencies responding to the survey, 10 reported implementing their evaluations in a way which develops government capacity. UNDP reported that all independent country programme evaluations include some elements which can contribute to the understanding of national counterparts of how an evaluation is carried out and its results communicated. Box 9 shows an example provided by UNICEF of learning-by-doing, using an evaluation to develop individual capacity. This example illustrates how capacity can be developed at the level of evaluation management (through engagement of the Government in the Evaluation Reference Group), and at the level of the evaluation team (team members receiving coaching and training from the lead evaluators). This should be a minimum level of national involvement and capacity development in any evaluation.

Box 9. Developing capacity in Timor Leste by involving the national counterpart in an evaluation of community-led total sanitation

A formative evaluation of community-led total sanitation was commissioned by the Timor-Leste Ministry of Health in collaboration with UNICEF Timor-Leste. The evaluation was conducted between December 2019 and March 2021, a period which included a hiatus from February–November 2020 due to the global COVID-19 pandemic. The evaluation benefited from the strong involvement of government and CSO stakeholders, as well as an inclusive approach at community level, with a focus on children, women and persons with disabilities. The Ministry played a key role in the development of the initial terms of reference for the evaluation, and was part of the Evaluation Reference Group responsible for quality control of the evaluation, checking whether its findings and conclusions were relevant, and recommendations implementable, and proposing improvements. To facilitate government ownership, and as the final report was written in English, the evaluation team had a face-to-face meeting with the Ministry to review the final report in the national language, Tetun. The Unit of Planning, Monitoring and Evaluation of the Prime Minister's Office was also part of the Evaluation Reference Group. The evaluation also contributed to building national evaluative capacity, as the evaluation team included ten Timorese researchers/enumerators (five women and five men, including three youth and one person with disabilities) who received training and regular coaching from the lead evaluators.

Source: UNICEF EAPRO.

As part of "learning by doing", several agencies reported involving government representatives in evaluation reference groups, which can contribute to institutional capacity development, as well as to individual capacity strengthening of those involved (see also section 3.4.3). UNFPA engaged young evaluators in the reference groups of country programme evaluations in Armenia, Georgia, Kazakhstan and North Macedonia.

Beyond involvement in evaluation governance mechanisms, a next level of collaboration is in doing the evaluation work together. FAO involved government M&E staff from the Ministry of Agriculture in survey analysis for its Tanzania country programme evaluation. It used government technical staff as team members in Indonesia, who participated in the field missions, data collection, evidence building and drafting inputs to the report. UNICEF is using the Kenyan National Bureau of Statistics for sampling and related work on one of their evaluations. The FAO Kenya country programme evaluation was done jointly with two national institutions under explicit agreements: one academic institution and one semi-public institution mandated to support the Parliament. Similarly, joint evaluations are a way to strengthen evaluation management skills on the job. WFP did joint evaluations in Benin, Dominican Republic, Eswatini, India, Lesotho and Namibia.

Another step up is to support country-led evaluations**,** as UNICEF did in Burundi and for an SDG evaluation on Social Protection in Madagascar. Mideplan in Costa Rica has started a process of capacity development for other institutions through the implementation of evaluations, which includes orientations, a collaborative design process including the development of terms of reference, implementation of the evaluation using participatory methodologies, and the review of recommendations eventually leading to management responses, learning and decision-making.

With respect to individual capacity, some agencies reported increasing the use of local evaluators. All of the case study countries appeared to have an adequate supply of local evaluators. For example, the percentage of local consultants hired by the FAO Office of Evaluation rose from 21 percent in 2014 to over 50 percent in 2020. The UNDP Independent Evaluation Office (IEO) emphasizes the use of local consultants, and since last year has actively sought to engage national think tanks or academic institutions to support components of evaluations, notably the analysis of the country context.

UNICEF shared learning on the danger of focusing on building the capacity of individuals, given the loss of impact when they move on, and the importance of focusing on systems and institutional strengthening.

3.5.2 Strengthening understanding of the role of evaluative evidence

The United Nations has an important role to play in building the understanding of decision makers of the role evaluation can play in policy and practice, as a precursor to generating demand for evaluation. This may involve advocacy, training and exposure visits.

Another aspect of individual capacity development is related to strengthening the enabling environment; building an understanding of the role of evaluations and evaluative evidence in policymaking and practice. This may precede, or be done in parallel with, more formal training.

United Nations agencies and other development partners can play an important advocacy role in strengthening understanding of the potential role for evaluation. In many countries the emphasis in M&E has been on the "M", or monitoring. In Kenya, this includes outcome monitoring for the annual performance review, and the midterm reviews of national development plans. Thus, in terms of monitoring and reporting Kenya is doing well, but formal evaluation has not featured strongly. In situations such as this, the potential of evaluation to unpack ***why*** things are working or not, and how they can be improved, is missing from the system.

Ten United Nations agencies indicated that they were supporting activities to strengthen understanding of evaluative evidence. In some cases, this was through training. For example, ILO reported a training programme which provided constituents with knowledge of the added value of evaluative evidence for policymaking,

accountability and reporting, with tools to effectively use evaluation for reporting progress on SDG 8 (decent work). UNIDO, UN Women and UNDP reported similar types of interventions. One respondent indicated that "we have to show policymakers how evaluation is useful in reaching policy goals. If government is convinced that evaluations help policies to be more effective, then ECD will start." To contrast the perception of evaluation as a "punitive" exercise, one respondent stressed the importance of moving beyond accountability and focusing on learning and building evidence of what works, how and why.

Other advocacy efforts include exposing decision makers to the experiences of others, for example through attendance at NEC conferences, or study visits to other countries. NEC conferences provide opportunities for evaluation champions to learn from and exchange with counterparts from around the world. Box 7 above describes the importance of a study visit for the development of the South African evaluation system.

Another approach to strengthening overall understanding of the potential of evaluation is to develop the capacity of evaluation champions, both individuals and M&E units, to play a 'knowledge broker' role. This is highlighted in a recent book on evidence use in Africa, which discusses the need to link those who demand evaluation with supply (Goldman & Pabari 2020). Knowledge broker courses have been developed to help participants strengthen skills such as: identifying the knowledge needs of policy actors; acquiring the necessary evidence; translating it into policy arguments; and reaching users with results at the right moment in the decision-making cycle.[21]

Other mechanisms used to strengthen the understanding of the role of evaluation mentioned by United Nations agencies are:

- Workshops (e.g. GEF Evaluation Sessions during the GEF Expanded Constituency Workshops);
- Seizing the opportunity of meetings and presentations with government officials to discuss evaluation and raise awareness, and being responsive to their interests (FAO);[22]
- Providing technical assistance, for example to evaluations related to national strategies of gender equality (UN Women), or seconding M&E staff to the Government (WFP);
- Ensuring that evaluation evidence is used in policy development (UN-Habitat, UNICEF).[23]

21 For example, the Knowledge Brokers Game-Based Workshop offered by Dominika Wojtowicz and Tomasz Kupiec at the 2019 NEC conference (https://nec.undp.org/workshop/2-knowledge-brokers-game-based-workshop)

22 In one example, a tailored video was produced in response to a demand for clarifications on evaluation raised by the Bureau Opérationnel de Suivi at the Présidence du Sénégal.

23 For example, the UN-Habitat research division produces evaluative evidence and a biennial flagship report to help Member States deliberate on their policies, and UNICEF works with CLEAR-AA to promote the use of evaluative evidence in VNRs in Africa.

3.5.3 South-South and triangular cooperation

Peer-to-peer learning between governments from different countries could be a powerful mechanism for national evaluation capacity development.

Peer learning between governments from different countries is increasingly promoted in other policy areas, and is an area where UNEG could play a role for NECD. In Africa, the 'Twende Mbele' programme brings together the Governments of Benin, Ghana, Kenya, Niger, Uganda and South Africa to share experiences and develop capacity and systems around M&E.[24] UN Women pointed to their work with EvalPartners and the EvalGender+ Initiative, a global multi-stakeholder partnership to promote equity-focused and gender-responsive evaluations. With UN Women providing technical support to regional-specific networks and national evaluation associations (including training, gender groups, and knowledge products), this approach has helped to build supportive environments for evaluation culture in different countries. FAO noted that communities of practice require dedicated attention, but can provide small and subtle ongoing capacity development which, in the long run, can make a useful contribution to NECD. WFP also pointed to the value of collaboration around international conferences (such as EvalMENA in the Middle East and North Africa, AfrEA in Africa, or the South Africa M&E Association) for several NECD activities in these regions.

3.6 Leading to change: demand, production and use of evaluations for decision-making

Core to the NECD theory of change is that public institutions demand evaluations. If they do, with the appropriate policies, systems and individual capacity in place to deliver and make evaluations publicly available, the stage is set for evaluation use, which it is hoped will ultimately lead to better policies and programmes, and better outcomes for people. This section examines progress on the demand and production of government-led evaluations, and United Nations support thereto.

3.6.1 Demand and production of evaluations by public institutions

Demand for evaluation depends on institutions and policymakers seeing the benefit of evaluation processes. It is therefore important to start producing evaluations in parallel with building national evaluation systems, to demonstrate their potential. In the case study countries, evaluations were still mostly commissioned by development partners, with a limited number of evaluations produced by government.

As mentioned above, at the heart of the NECD theory of change is demand for evaluation. Costa Rica is the only country that has developed a formal national evaluation agenda. In the Philippines and Morocco, while there is demand for evaluation, this has not yet been fully expressed through a national evaluation agenda.

24 www.twendembele.org/

© Omar Awabdeh

FAO country programme evaluation in Indonesia conducted jointly with government technical staff, 2019

There is not currently much evidence of demand in Kenya, while in Sri Lanka, demand existed when there was a functioning system in the 2000s, but is less evident today. In most case study countries, with the exception of Costa Rica, demand for evaluation still comes mainly from development partners.

Many governments or government institutions remain wary of evaluation, fearing that it is punitive or not recognizing the potential benefits for performance improvement. Only by producing evaluations that are seen as useful will policymakers be convinced that this is a process worth supporting. For this reason, it is key to start conducting evaluations early, and not wait until the complete system is established. Rather, by carrying out evaluations and demonstrating their utility, policymakers will be more motivated to further invest in their national evaluation systems. Some of the case study countries started work on developing systems without actually undertaking evaluations (e.g. Kenya, Sri Lanka), while others carried out evaluations in parallel with system development, but in a fragmented way (e.g. the Philippines).

In carrying out the case studies, it proved difficult to get a full picture of the production of evaluations by the public sector. **Table 7** shows the data available from the case studies, which suggests that evaluations are being produced in these countries, but the number of government-led evaluations remains significantly lower than donor-led evaluations. The opposite may be true in middle-income countries such as Mexico and South Africa. As for United Nations support, six agencies indicated that they had supported government-led evaluations (UNDP, UN Women, UNFPA, UN-Habitat, UNICEF and WFP).

Table 7. Evaluations produced by government and available on public databases

COUNTRY AGENCIES	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica	Other examples
Government evaluations	NEDA has undertaken some through UNDP and 3ie	None recently	3 pilots in 2017	21 since 2010	7-8 impact evaluations per annum (ONDH)	15 from 2015-2018 in the NAE 60 planned for 2019-2022[25]	South Africa 62 national evaluations 2011-2017 (plus provincial and dept evaluations)
United Nations agency evaluations since 2019	32	36	43	25	4	2	South Africa 34
United Nations agency evaluations since 2011	137	72	172	76	10	14	South Africa 98
3ie evidence portal, impact evaluations since 2011	40	59	195	6	32	8	South Africa 306 since 2011 Mexico 68 since 2011
Other	DFAT/3ie 5 IEs		276 from 2005[26]				

In Benin, Costa Rica, Morocco and the Philippines, the case studies show efforts by governments to commission or undertake evaluations, with a number of evaluations being produced. In Sri Lanka, evaluations were being undertaken in the 2000s, but not currently, and in Kenya it is not yet happening, although some of the systems to enable production are in place. Costa Rica has a National Evaluation Agenda, with 15 evaluations completed from 2015–2018 and 60 evaluations planned over the Agenda period. Benin has conducted 21 national evaluations in the last ten years. In Morocco and Kenya, there is also emerging interest from local governments.

25 38 percent being done internally to reduce costs and develop capacity.

26 CREST/CLEAR AfreD database.

The lesson emerging is that it is not advisable to work on developing national evaluation systems without starting on producing evaluations (as happened in Kenya or Sri Lanka). Advocacy around evaluations is much easier if there are evaluations to demonstrate.

3.6.2 The use of evaluative evidence in decision-making

The intended outcome of support to NECD is that countries use evaluative evidence to inform decision-making. The survey and country case studies provide anecdotal evidence of the use of evaluations in decision-making, but in most cases there are no tracking mechanisms to systematically monitor this.

Seven responding agencies felt that their evaluations had contributed to government decision-making. GEF felt that most of their evaluations did, while UNICEF, UN-Habitat, UNFPA and WFP reported that some of their evaluations contributed, and ILO and UNIDO reported a few. Only UNICEF, ILO and GEF reported systematically monitoring the implementation of recommendations.

A UNICEF respondent shared that: "Tangible examples of the actual use of evaluations in decision-making remain hard to find. Feedback and reporting from country offices at times provide examples but these can often be anecdotal. It can take years for the sought policy, practices or budgetary changes to come into effect."

Only three agencies reported some system for monitoring use (GEF, UNICEF and ILO). ILO EVAL's automated management response system looks at recommendations addressed to governments and asks the country office to provide updates on progress in their implementation. The GEF Independent Evaluation Office conducts stakeholder needs surveys. UNICEF Tanzania indicated that they reviewed key planning documents to ensure that they had adequate references to evaluative evidence. They also asked sectoral colleagues for regular discussions with national government counterparts and updates on their responses to evaluation recommendations. UNICEF also has a management response system, and the 'Influential Evaluations Initiative' asks regions to provide case studies of evaluations that have proven to have led to positive changes.

From the country case studies, Benin's BEPPAAG undertook follow-up research on the implementation of recommendations from evaluations carried out between 2010 and 2013. The report found 90 recommendations from nine PPEs, 70 of which were included in plans by the concerned line ministries. Of these 70 recommendations, 39 (56 percent) were fully implemented at the time of the follow-up mission. Of these 39 recommendations, 19 led to the review and/ or development of new public policies (Primature, 2016).

Several examples of use were reported in the case studies and interviews. One excellent example was the UNICEF-supported evaluation of the 'Thailand Cash Grant' programme, which led to additional funding from the Government. WFP reported five examples of use of evaluation results including in Peru where the agency provided technical assistance to an impact evaluation of the national school feeding programme that was managed

by the Government. The results of that evaluation fed into re-design of the programme (EvalForward 2021). In Tunisia, evaluation recommendations on WFP capacity strengthening activities for the national school feeding programme were used for decision-making in scaling up the project (the budget was doubled). In the Philippines, Congress expanded the coverage of the conditional cash transfer programme to cover the children of poor families attending secondary school, citing evaluation results which showed that the improvement in enrolment rates among children of poor families as a result of the programme stopped at the primary level (CPRM, 2019).

Promoting the use of evaluative evidence requires a strong focus. The motivation of policymakers is key, as one respondent indicated: "Interest and demand is the main thing, if people are competent and motivated, they can push in their institutions." United Nations agencies can support national government and non-governmental actors to create the motivation, opportunity and capacity for policymakers to use evaluative evidence.

3.6.3 Overall progress of the case study countries

Evidence suggests that there has indeed been progress in NECD since the passage of the General Assembly Resolution in 2014. The progress in case study countries cannot be linked directly to the Resolution, as most people interviewed had not heard of it, but could have an indirect link through enhanced engagement by United Nations agencies.

The case study sample included countries with evaluation systems which have emerged since 2014 (Kenya, Philippines), which were already established (Benin, Costa Rica, Morocco), and one which had advanced in the 2000s but then fallen back and has been moving forward again since 2014 (Sri Lanka). A summary of the situation of these countries, as assessed against the short- and medium- term outcomes of the theory of change, is provided in **Table 8**. This summary indicates that Costa Rica has the strongest systems in place, followed by Benin, Morocco and the Philippines.[27]

Costa Rica (an upper middle-income country) and Benin (a low-income country, reclassified by the World Bank in 2020 as a lower middle-income country) have continued a steady progression. Morocco has continued with a strong independent government agency (ONDH), but a system that is not government-wide. The Philippines has moved forward, although not as much as might have been expected once the policy framework was approved in 2015. Sri Lanka has continued to be active in the VOPE space, and from 2018 there has been some progress in the Government and Parliament. Kenya has also seen some progress with an active VOPE and has guidelines since 2019–20. In the latter two cases government evaluations are not yet being commissioned. All the case studies, except for Costa Rica, have been supported by UNICEF and UNDP, and a wide range of agencies have provided some support in each country.

27 A forthcoming chapter of a handbook on public policy evaluation reviews progress across Latin America, Africa and Asia and will give a fuller picture of the state of national evaluation systems in lower middle-income countries.

Table 8. Overview of the theory of change outcomes

Outcome	Philippines	Sri Lanka	Kenya	Benin	Morocco	Costa Rica
Impact on policymakers						
Use of evaluative evidence in decision-making	Some evaluations used.	No.	No.	Survey indicated yes.	Some evidence.	Used in planning and budget.
Medium-term outcomes						
Production of quality evaluations by public institutions and made publicly available	Produced, but not systematically.	Not being produced.	Not being produced.	Produced systematically.	Produced systematically.	Produced systematically (plan).
Production of quality evaluations by non-public institutions and made publicly available	Unknown[28]					
Short-term outcomes						
Enabling environment						
Policy and regulatory environment enables evaluation	Policy in place not law. Draft bill.	Policy in place not law. Draft bill.	Policy in place not law.	Policy in place not law. Law being prepared.	Constitution refers to evaluation, decrees not policy.	Law and policy in place.
Stakeholders exchange knowledge and coordinate around the national evaluation agenda	In government. Limited externally. Conferences.	In the VOPE. Evaluation weeks.	In the VOPE. Some interaction with government in evaluation weeks.	National Evaluation Council. 2 yearly evaluation weeks.	Yes through ONDH. Not a national agenda.	Participate in National Eval. Platform. Unis very active.
Institutional capacity						
Public institutions demand evaluations	Some demand – no plan.	No demand.	No demand. Plan being prepared.	Demand.	Demand.	Demand and national eval plan.
Evaluation champions in centre of government and more widely	NEDA (DBM), some M&E units, Parliament, 2 VOPEs.	DPMM, some M&E units, Parliament, Very active VOPE.	MED, some M&E units, very active VOPE.	BEPPAAG very active, VOPE.	ONDH very active, VOPE.	Mideplan active, some M&E units, university, 2 VOPEs.
Appropriate institutional systems for evaluation in place (structures, frameworks, human resources, budgets)	Guidelines, some M&E units.	Policy Framework with some guidance.	National/ subnational guidelines.	Guidelines, Repository, 60% funded by gov.	Part-funding evaluations.	Guidelines, self-funded evaluations.
Non-public sector organizations demand evaluations and offer ECD services	Demand unknown. Services in all countries include provision of evaluators and training.					
Individual capacity						
Individual capacity to manage, conduct and use quality evaluations	Some capacity. No post grad M&E.	Some capacity. Post grad diploma.	Some capacity. Masters M&E.	Some capacity. Masters M&E.	Some capacity. Masters M&E.	Some capacity. Masters M&E

28 The short time to do the case studies did not provide enough time to follow-up on this aspect.

4. Reflections on the way forward for UNEG

This section builds on the discussion in section 3 to draw lessons and propose reflections for UNEG on its roles and responsibilities for NECD. It argues that there is potential for United Nations agencies to carry out their own evaluations with a stronger NECD perspective, and to increasingly support joint and country-led evaluations. Following this, the section shares lessons emerging from the case studies and other experiences on direct support to NECD, insights on the engagement required in terms of policy or leadership decisions and resources, and the importance of strengthening coordination and collaboration for greater impact. The section closes with reflections on the importance of NECD for the SDGs, the role the United Nations can play, and on the pressing need to ensure that national evaluation systems are able to integrate critical issues such as climate change and inequalities into their work.

4.1 Undertaking United Nations evaluations with a NECD perspective

4.1.1 Strengthening meaningful national engagement in United Nations evaluations

There is potential for United Nations agencies to reconsider the way they conduct evaluations. To strengthen national capacity while meeting agency evaluation requirements, United Nations agencies can rely more on local experts, engage government counterparts more meaningfully in evaluation governance mechanisms, use national evaluation systems or tools once in place, and promote truly joint or country-led evaluations.

As discussed in section 3.5.1, United Nations agencies can implement their evaluations in ways that foster learning-by-doing to build individual capacity of both government counterparts and evaluation consultants. Since the beginning of the COVID-19

pandemic, agencies have increased their use of local evaluators, reinforcing the lesson that local evaluators bring a vital understanding of local context in addition to other expertise. Local evaluators should be seen as more than just data collectors. The case studies and literature review show that many countries have a good pool of qualified, national evaluators. Even in countries where there are fewer evaluation experts, there are likely to be subject matter experts or emerging evaluators who can be paired with international evaluation experts, whose terms of reference can be expanded to include mentoring local team members. For example, UNFPA included young and emerging evaluators in the evaluation team for a regional programme evaluation of the Arab States, a practice that could be expanded.

Beyond evaluators, there is a lot of potential for United Nations agencies to contribute to NECD through enhanced national ownership of United Nations-led evaluation processes, by promoting the meaningful engagement of government and non-governmental partners, for example, in steering committees, reference groups and other forums. Wherever possible, government partners should be encouraged to take a leadership role in these mechanisms.

This collaboration could be expanded by encouraging genuine joint and country-led evaluations. It emerged in the course of the study that the term "country-led" evaluation is sometimes misused. In the United Nations context, some use the term to describe situations where the country (government) has some involvement in a governance mechanism of a United Nations-led evaluation (such as a reference group). The term "country-led" should refer to situations where the country (government) takes a clear leading role in planning and commissioning an evaluation that responds to its own priorities, with the United Nations playing a supporting role.

One government respondent commented that the United Nations "…don't walk the talk – they talk very nicely when they are planning, e.g. when they do the UNSDCF, they invite the Government to chair, but during implementation they go their own way. The policy and incentives that drive them come from their headquarters and the interest of the staff is to impress their bosses, to deliver on their outputs." In the case of evaluation, in most cases United Nations agency evaluation offices have been established for internal accountability and learning or accountability to donors. Introducing a focus on NECD would require a change in perspective and a long-term vision.

Another step would be to implement United Nations evaluations through the government system, thus helping to strengthen that system and create greater awareness through the process. A genuine country-led partnership can include joint funding, particularly for middle-income countries, as seen in the example of the 2009 South African UNDAF evaluation (Box 8). This may lead to interesting results. For example, in the South Africa UNDAF case, the South African Government led the evaluation of United Nations work in the country, and the final report was quite critical in some areas. This is reported to have led to some "soul searching" in all United Nations offices, and was used to guide United Nations work in the country for some time.

There may be challenges in promoting these joint, capacity development-oriented approaches. National counterparts may not be ready for a potentially critical assessment, and evaluations may touch on sensitive issues that national counterparts may not wish to bring to light. There may be differences in perspective between the United Nations and the government on particular issues (e.g. gender equality, human rights, etc.) As one respondent said: "Evaluation can always be used as a weapon at national level, for government or the opposition, or there can be a serious fight if the findings are not favourable. Evaluation is a serious tool that has to be managed tactfully." For this reason, it is important to have inclusive processes and base dialogue on strong evidence.

It should also be recognized that the United Nations may have its own requirements that mean it needs to carry out more evaluations that the national system can absorb. Some steps that United Nations agencies can consider to prioritize evaluations to be carried out with an NECD perspective include:

- Deciding which evaluations are motivated more by internal United Nations interest, and which are of significant interest to partner governments. In the latter case at least, governments should be contributing to the terms of reference, leading reference groups, and participating in technical working groups;
- Identifying which evaluations are clearly linked to national priorities and could be on the national evaluation agenda, though financed by the United Nations agency;
- Identifying where United Nations agencies can use elements of country systems, like templates for terms of reference, guidelines or management responses. United Nations agencies may have helped to develop these systems.

4.1.2 Adapting time frames and processes to enhance participation and capacity development

Conducting evaluations using an NECD approach may require more time to create space for learning and capacity development. United Nations agencies need to balance long-term objectives with short-term requirements, seizing opportunities for NECD and showing a willingness to depart from potentially rigid processes. The potential for two-way learning needs to be recognized: joint evaluations are learning opportunities for all partners involved, including United Nations agencies.

Joint evaluations with government contribute to strengthening partnership and ownership over the evaluation at national level. They provide a capacity development opportunity not only for countries but also for United Nations agencies, as one respondent highlighted: "Joint missions (with countries) ha ve been transformative, opening up space for dialogue and thinking."

Undertaking evaluations with a capacity development approach is likely to require more time and a slower pace to create opportunities for more joint work and consultation and allow for strengthening capacity. A United Nations agency evaluation specialist

shared that: "In some countries in South East Asia where capacity of stakeholders on (monitoring and) evaluation is limited, if you seriously engaged people more in management (including decision-making) and in conducting evaluations, the evaluation process certainly will take a longer time … but capacity and understanding will genuinely be improved. You need to be patient and be prepared to experiment and see. But how to involve them to take it seriously and how to … engage them … in the evaluation process (not just inform or consult them)? Are we ready for 'capacity development' by working with them more seriously and let go of certain requirements of 'high' quality evaluation reports that meet a quality standards checklist...?"

Another United Nations respondent stressed the importance of a long-term vision and readiness to respond to emerging interest, even with little support, commenting: "To make meaningful progress, NECD initiatives often need a combination of playing the long game, nurturing and supporting champions, identifying where commitments and interest are there or latent, allocating resources at critical moments to inspire action and show the benefits of good quality evaluation and evidence generation, and communicating very smartly about this. Balance this with investments in "emergent" countries – at times the level of appetite is surprising and this needs to be kept enthused – even if (this means only) a small-scale investment in the short term."

Another approach to working in a way that builds capacity is to use evaluation consultants as facilitators, rather than just outsourcing evaluations to them. For example, Twende Mbele is supporting training in rapid evaluation in different African countries. One of the advantages of the rapid evaluation model used by Twende Mbele is that it is possible to consider using internal staff, which reduces costs and builds capacity. Three options are proposed in these rapid evaluations: internal, outsourced or 'facilitated'. In the 'facilitated' option, external expert evaluators support an internal team to carry out the work (South Africa DPME 2020).

One WFP respondent raised an important point: evaluators who engage in NECD need to develop their facilitation skills and think beyond the technical aspects of evaluation. They need to pay attention to how to make evaluation relevant to national policymaking processes. The WFP secondee to India's DMEO demonstrates another way to use consultants in a developmental way.

4.1.3 United Nations Sustainable Development Cooperation Framework evaluations and NECD

UNSDCF evaluations provide great opportunities to use national systems, where capacity is in place.

Many of the points discussed above apply to system-wide, and in particular UNSCDF, evaluations. UNSDCF evaluations are good candidates for the use of national systems, as these frameworks are intended to support national government priorities and are linked to the SDGs. Where a national system is functioning, national governments could lead these evaluations to assess the overall contribution of the United Nations in the

country, which would have a big capacity development effect, as was the case in South Africa. However, this does pose challenges: these evaluations are decentralized and as such there are some credibility risks in the country, so a robust, credible evaluation is needed. Advocacy and support may also be required to convince Resident Coordinators and agency heads in-country of the advantages of this approach.

4.2 Direct support to strengthening national evaluation capacity and systems

4.2.1 Lessons from the case studies on supporting national evaluation systems

The process of building a national evaluation system needs to be seen as a long-term project (over 10-20 years), with medium-term perspectives (over five years), during which a lot can be achieved. It is as a continual, gradual process. A key message is the importance of starting to conduct evaluations to demonstrate their potential, alongside developing the system, and thinking about how the evaluation process can strengthen the evaluation system.

Several lessons emerge from the country case studies that can inform United Nations agencies on how to move forward with supporting national evaluation systems. The first lesson is that developing a national evaluation system is a long-term process, and needs to be gradual to allow for more sustainable entrenchment of evaluation culture and practice. The Benin case study illustrates this, as the national evaluation system was started under the leadership of the Government, but from the outset, bilateral (Belgian, Danish, Dutch, French and German) and multilateral (UNDP, UNICEF and World Bank) partners played a significant role in supporting the evaluation system, support which has continued for over ten years. One respondent, referring to the follow-up to the United Nations Resolution stated: "if we really want to push for change, we need to be consistent and continue to follow up, there is a need for a long-term vision and build incrementally to that vision."

It may also take time to build trust in partners or in the concept of evaluation. UNODC found it took two to three years of doing small activities in Morocco to build trust, after which they were able to scale up over time. They realized that NECD is a long-term investment: "it takes a long time and involves learning-by-doing but it can be very rewarding to see changes happen." The process can also start with the promotion of 'evaluative thinking'. For example, WFP conducted a webinar on evaluative thinking with country offices. For countries with strong monitoring systems, but limited interest in evaluation this is a good way to stimulate interest, and should be part of an effective management system.

Some case study countries provide evidence of the development of an evaluation system without evaluations actually being implemented (e.g. Kenya and Sri Lanka). Other countries, such as Benin, Costa Rica and South Africa, have developed elements

of the system incrementally while they take forward evaluations (Goldman ***et al***. 2018). Thus, as previously mentioned, it is important to start doing evaluations and use these to advocate for doing more. This was highlighted in the 2012 UNEG Guidelines which suggest to: "Work with country officials to build into the strategy the conduct of an evaluation early on in NECD, as a way to demonstrate to senior officials the utility of evaluation as a tool. Work with the central unit to identify a topic that is of importance to senior officials but not overly ambitious, so that there is a high probability that it can be successfully evaluated and reported within a year." (UNEG, 2012).

United Nations agencies could play a key role in supporting these early evaluations, though with a broader perspective than the evaluation itself. One respondent from the Global Evaluation Initiative (GEI) suggested: "Many think that doing evaluations is a way of building capacity, but what we need to think about is how an evaluation strengthens country systems. In practice, for example, who needs to actually do the evaluation, how are the recommendations managed, how is the resourcing provided, what are the procurement arrangements, who designed the key evaluation questions, how does that process build capacity."

The UNEG Guidelines also suggest that: "In advising countries on how best to fill a void (when evaluation capacity is non-existent), don't recommend an immediate whole-of-government uptake, but rather a more selective piloting, where evaluation can be phased in over an appropriate number of years (say five years, starting with the strongest ministries as pilots). This could allow for a period of learning and adjustment, as well as evaluation of human resources skill building. A government-wide roll-out at one point in time would probably over-extend the resources of a government's central unit that is facilitating M&E development." (UNEG, 2012).

Another way of building confidence can be to undertake rapid evaluative exercises which can quickly feed back into policy processes, and overcome a perception that evaluations are long and costly. The COVID-19 pandemic has led to a demand for rapid processes, and countries like Benin, Ghana, Kenya and South Africa are undertaking rapid evaluations with the support of the Twende Mbele programme.

Another very useful entry point can be to systematically search for evaluations conducted in the country, and develop an evaluation map which can be drawn on by policymakers. This shows which evaluations exist, but also where there are gaps where new evaluations would be beneficial. Uganda has done this and found over 600 existing evaluations, which can be used immediately to inform policy and practice.[29] This also creates a repository of evaluations.

With respect to evaluation system development, a key early stage is to bring together national stakeholders to build a common understanding of what sort of system is desired, and start to develop the effective collaboration of all stakeholders to develop the ecosystem within which evaluation can thrive. In South Africa, the group who went

29 https://chs.mak.ac.ug/afcen/map/ [Cited 30 July 2021].

on the study tour to Mexico and Colombia formed the nucleus of a national Evaluation Technical Working Group, which supported the DPME to develop the system (Box 7). In Costa Rica, the two-year participatory and collaborative process of developing and approving the NEP has contributed to reinforcing the legitimacy of Mideplan and the effectiveness of its major coordination instrument: the National Evaluation Platform.

For some countries, the early development of a policy is critical. Respondents from Sri Lanka commented that without a policy it is difficult to get agreement from all the stakeholders. In other cases, starting evaluations was more important. It is important to at least develop some core principles (e.g. a utilization focus) to inform how evaluations should be implemented, and then start doing evaluations.

From the United Nations perspective, UN Women stressed the importance of continuity, sharing successful examples from Colombia and Georgia where evaluation specialists had been in post for over five years and managed to build up an understanding of the context and relationships to build the system.

4.2.2 Entry points for wider NECD work

There are multiple entry points for United Nations agencies to engage in NECD, which include SDG-based policy planning and implementation, the preparation of VNRs, support to national development planning and budgetary processes, advocacy with top leadership and parliament, assessments of M&E systems and capacity, and funding of priority evaluations.

There are a wide range of entry points for strengthening evaluation capacity development by United Nations agencies, beyond the work of the evaluation units themselves. This includes support to SDG-based policy planning and implementation, and/or to strengthening SDG monitoring, which could be augmented with support to SDG-focused evaluation. Similarly, support to national development planning and budgetary processes can be further strengthened with advocacy for, and support to, capacity development for evidence-informed planning, with evaluation as part of these processes.

Another key entry point is support to the preparation of VNR reports, where it is possible to advocate for the inclusion of evaluative evidence in the reporting, as well as support to develop that evidence base (see for example the work by UNICEF/CLEAR AA in Africa). Nigeria used 46 evaluations to substantiate its analysis for its 2020 VNR, and Madagascar did likewise for its 2021 VNR. This also serves to demonstrate the utility of evaluation.

Advocacy with high-level leadership and parliaments to take evaluation seriously is another good starting point. This could include peer-to-peer visits with countries such as Benin, Chile, Colombia, Costa Rica, Mexico, South Africa or Uganda, or training of senior managers in evidence-based policy, as supported by Twende Mbele.

M&E system assessments can create space for dialogue on what is already in place and what could be strengthened. United Nations agencies have already conducted assessments of evaluation capacity, and in Latin America supported the development of a national evaluation capacity index. ILO and UNDP are using participatory evaluability assessments to identify opportunities to implement national capacity development programmes on the SDGs. GEI has developed an M&E System Analysis tool, drawing in part on the experiences of United Nations initiatives, which is designed to lead to ECD strategies, the implementation of which can be supported by different agencies.

Funding of priority evaluations can also spark interest, and every evaluation is an opportunity to build evaluation systems and capacity.

4.3 United Nations agency engagement in NECD and resources

National evaluation systems development and NECD is complex and requires a long-term vision and consistent support. Ideally, NECD should be reflected in all United Nations agency policies, and become a way of working. Direct support to national evaluation systems and NECD requires dedicated budgets, of individual agencies or joint programmes.

National evaluation systems and NECD are complex. If the United Nations is to take forward NECD effectively, agencies need to be consistent, and build incrementally towards a stable, long-term vision. This should be expressed in agency evaluation policies, so that directors become accountable for taking NECD forward. NECD also needs to become a recognized element of work, seen by executive boards as a key priority.

Policy is not the only determinant for NECD. A lot also depends on the management of evaluation units**,** which can prioritize support to NECD, at a minimum by adopting a capacity development approach to the conduct of evaluations, even if NECD is not explicitly mentioned in their policies.

NECD needs its own budget, so that it is not just an “add on”. NECD needs to be seen as a long-term project, with medium-term commitments (at least four- to five-year terms) to consolidate national evaluation systems. This could be budgeted for within individual agencies or through joint programmes. United Nations agency funding can leverage national resources, as seen in Benin where around 60 percent of funding for country-led evaluations is now coming from the Government.

In terms of funding, one agency reported that NECD is not seen as high priority for programme staff in country-based positions, which links to the limited resources to support NECD work. Another agency respondent reported that many donors consider NECD, along with M&E, to be an administrative task and cost, and do not want to fund it, although it requires a long-term commitment. NECD needs to be seen as part of

governance, and important for governments to be able to build up a knowledge base, learn and be accountable.

Finally, adapting United Nations evaluation processes to incorporate a NECD approach does not necessarily require additional financial resources, although it may require additional investment by staff.

4.4 Coordination and collaboration

4.4.1 Coordination and collaboration among United Nations agencies

United Nations operations can be fragmented, and there are many barriers and lack of incentives for collaboration which can weaken, rather than strengthen, national capacity. There is scope to strengthen collaboration and synergy, within the United Nations and beyond at country, regional and global levels. NECD is a good candidate for joint programming, and can also become part of United Nations country team M&E group agendas. It should be seen as part of a wider strategy of bringing learning for continuous improvement into government.

Respondents to the survey pointed out that United Nations work on NECD is often fragmented, and that there can be many barriers and lack of incentives to collaborate. Some of the challenges mentioned are internal to the United Nations, for example, different sectoral foci, different governance systems and requirements, competition for resources, mixed signals from leadership, as well as differences between agency planning and evaluation offices/ units which often have different government counterparts. Other challenges may be external, as governments should ideally take the lead on coordination to ensure that they receive the necessary support for their priorities, but this does not always happen for a range of reasons.

United Nations reform includes a call for more joint programming,[30] and NECD is a strong candidate for joint action. The case studies highlight the importance of coordination, particularly between UNDP and UNICEF who are often working on different aspects of NECD in the same country. In the Philippines, where UNICEF and UNDP were both working on NECD, there was no joint programme and according to respondents there was a degree of competitiveness. The Morocco case study illustrates the potential for joint programming, as described in Box 10, particularly relevant in this middle-income country with a relatively small United Nations presence.

30 See for example, UN General Assembly Resolution A/RES/67/226, Quadrennial comprehensive policy review of operational activities for development of the United Nations System, adopted on 21 December 2012, para 18, which "Encourages the United Nations Development System to further strengthen joint programming processes at the country level, where appropriate, as a useful way to promote greater coherence, taking into account the principles of national ownership, alignment with national priorities and the comparative advantage of individual entities of the United Nations system at the country level."

Box 10. Joint programming in Morocco

Since the majority of public policies are multisectoral, and to maximize the United Nations contribution to the country, the United Nations system has developed two joint programme instruments to reinforce NECD in Morocco. First, support to ONDH in the framework of two phases of a joint programme that began in 2013, implemented by UNDP, UNICEF, UNFPA and UN Women, and temporarily supported also by UNODC. Activities carried out by the programme included ONDH staff training, developing regional workshops, studies, surveys and other realizations designed to produce reliable data and knowledge for the analysis and evaluation of public policies. This has contributed to build ONDH capacity for design and use of evaluation, new tools and methodologies. This programme was evaluated in 2016.

The second instrument is the Joint Programme to Strengthen Harmonization and Public Policy Evaluation (HEPP), in collaboration with UNDP, UNICEF and UN Women. It was coordinated by the Ministry of General Affairs and Governance from 2013, and was originally intended to support acceleration of the Millennium Development Goals by strengthening governance and public policy effectiveness on the basis of effective coordination and evaluation. The second phase (2017–2021) includes workshops on social protection, public policy alignment in the sector of public health, a regulatory framework for evaluation, guidance and methodologies for PPE, South-South cooperation and formative exchanges between public officers.

The two programmes align with priorities in the UNDP CPD 2017–2021 to bring coherence between national and regional development plans, support the Government in meeting commitments on national evaluation capacity,, develop transparent and results-based M&E, and partner with universities, Parliament and other institutions to further institutionalize the M&E culture, particularly targeting local governments (UNDP 2016).

The United Nations country team M&E Working Group has the responsibility of monitoring and evaluating the UNDAF. It is also responsible for sustaining the process of building national capacity in evaluation of the United Nations staff and its counterparts. In its function of quality assessment, it is responsible for providing support in methodologies and data collection and providing inputs for communication on M&E. The M&E Working Group is composed by a member of the country team (currently UNICEF) and includes a member of the Government (currently ONDH) (United Nations 2011).

As regards United Nations work at the country level, NECD should also be included in the cooperation framework. However, United Nations respondents explained that the cooperation framework is usually developed by programme teams who may not consider M&E themes. Where NECD is included, it is often only a small component of a broader governance programming framework. NECD should be seen as part of a wider strategy of bringing learning for continuous improvement into government. Senior-level engagement and interest can help to ensure that this cross-cutting theme is embedded in the cooperation framework and leads to coordinated interventions.

National Evaluation Capacities (NEC) Conference 2017, Turkey

NECD should also become a regular item on the agenda of the United Nations country team M&E meetings as a cross-cutting topic, which is potentially relevant for all United Nations agencies present in the country, particularly with respect to how they carry out their evaluations. As described in Box 10, in Morocco, the country team M&E Working Group is not only responsible for the M&E of the United Nations cooperation framework (its typical function) but also for supporting the process of building national evaluation capacity, and includes in its membership a representative from the Government. The Philippines and Sri Lanka show that despite the need for coordination among agencies, in country team the M&E Working Groups there is little discussion on NECD and no joint planning.

Coordination at country level needs to go beyond United Nations mechanisms. The UNEG NECD Practical Tips guide states that: "the national champion for M&E (for example, a representative of the central agency leading NECD initiatives) needs to be considered as an equal partner and participate in all such discussions (around coordination of NECD initiatives)" (UNEG, 2012). In countries with a national evaluation structure, the United Nations should engage with that mechanism. For example, in Costa Rica there is a National Evaluation Platform, to which the United Nations has been invited, but has not participated. According to respondents for the Costa Rica case study, donors and agencies do not share common frameworks, methodologies or approaches to build evaluation capacity, in spite of the country's success in developing an evaluation system.

There is also potential for greater coordination at the regional level to strengthen support to NECD. The strongest United Nations evaluation coordination forum is in the Asia-Pacific region, the United Nations Evaluation Development Group for Asia and the Pacific (UNEDAP). This was established over 10 years ago, and the 11 regional M&E advisers are very committed and co-organize training on ECD for agencies at country

level. They support country offices in a coordinated manner, with joint missions, running training for agency staff and providing support to UNDAF/ UNSDCF evaluations, with specific agencies delegated to support specific countries. A similar structure is being established in Eastern and Southern Africa. However, UNEDAP does not necessarily support NECD. There is potential to further strengthen this good model by adding a NECD lens to the work.

Global coordination could also be further strengthened. UNEG is already a global coordination forum for evaluation. There is scope to further link the work of different UNEG working groups. For example, United Nations reforms with respect to evaluation have focused on "system-wide evaluation", but this remains internal to the United Nations. There is scope to bring in elements of NECD into system-wide evaluations (including UNDAF and UNSDCF evaluations), as discussed earlier.

4.4.2 Coordination and collaboration beyond United Nations agencies

There is potential for collaboration beyond UNEG to strengthen NECD at country level and globally, notably with the GEI and EvalPartners.

Many other actors, beyond the United Nations, support the strengthening of national evaluation capacity and systems, thus coordination and collaboration need to extend beyond the United Nations family. There are already good examples of United Nations collaboration with other actors, such as the collaboration between WFP and DEval to develop the INCE in Latin America. From the outset, the INCE was conceived as a collective creation, involving representatives from government evaluation units, professional evaluation networks, academic institutions, CSOs and international development partners. The objective of INCE is to measure evaluation capacity and practices in the fields of policy, programmes and social services, as input into a periodic report that summarizes the characteristics and trends observed in each country. This is expected to support the further development of national evaluation capacity in partnering countries. In Sri Lanka, where several other agencies are supporting evaluation as part of the good governance agenda, some respondents noted that the country would benefit from these activities happening in a more coordinated manner with less competitiveness, building on each other's achievements and covering the gaps (Trikawalagoda, 2019).

One of the older initiatives promoting collaboration is EvalPartners, along with its offshoots EvalGender, EvalSDGs, etc. EvalPartners has been very successful in raising the profile of evaluation, for example, with young emerging evaluators and parliaments, and in communicating the link between evaluation and the SDGs.

A new initiative is GEI, which aims to support countries to strengthen evaluation and monitoring systems and capacity. Partners include the World Bank, UNDP, the World Bank-founded CLEAR Initiative with its network of seven regional centres, WFP, UNFPA, DEval and many others. GEI was created in part due to the recognition of the fragmentation of ECD support. Thus, part of its mission seeks to catalyse strategic

partnerships, bringing together key actors and experts in the evaluation field to help governments place evidence at the heart of decision-making. GEI targets three levels of support for ECD to address the M&E needs within a country, which are closely related to the UNEG theory of change:[31]

- Supporting the *enabling environment*, focusing on strengthening the legal, regulatory and institutional frameworks that promote a culture of the use of evidence in decision-making and lead to system-wide impact.
- Supporting *institutional capacity*, to build the M&E frameworks and capacity of institutions.
- Supporting the *capacity of individuals*, whose knowledge, skills and competencies are essential to achieving any system-level impact or culture change.

The GEI model advocates for the development of an M&E capacity development strategy at country level, which must be owned by the country's government and can be supported by local and international partners. According to the GEI Programme Manager, this: "should draw on evaluation and monitoring, but also on other forms of data research and statistics, and it must be use- and solution- oriented. What we would want to see is the growth of learning organizations that are accountable to their stakeholders … a systematic collaboration convened by a country authority and supported or facilitated or enabled by a GEI or related partner. The idea of the partnership is to help work out who is doing what, what are the needs of the country over the next five to ten years, and to help design programmes to deliver results." In other words, coordinated support to NECD is more likely to lead to sustainable results.

What does this mean for the role of UNEG? There is potential to use GEI at country level to systematically support national efforts from an evaluation perspective, and encourage United Nations agencies and other relevant development partners to collaborate in a mutually reinforcing way, mobilizing a pool of resources to support a common, nationally-owned NECD work plan. This can have a multiplier effect as illustrated by the Benin and Uganda examples mentioned earlier.

Respondents suggested some advantages of collaborating with GEI for UNEG. Firstly, as NECD is not an agenda that resonates with all UNEG members, especially the smaller agencies, United Nations agency evaluation offices and staff are not necessarily strong in capacity development. As such, GEI and its partner CLEAR Centres may be better positioned to think about evaluation from a policymaker's perspective, and bring in the necessary technical expertise to support institutional capacity development. GEI could also help to create a framework for NECD, creating synergies and systems from different initiatives at country level. GEI provides a space slightly outside the United Nations, to which United Nations agencies and UNEG bring specific technical and sectoral expertise which can benefit partnerships at country and global levels. Some respondents also pointed to concerns about World Bank

31 https://www.globalevaluationinitiative.org/our-work [Cited 30 July 2021].

dominance in the GEI initiative. Others highlighted the importance of United Nations specificities, as United Nations agencies share a set of values, a way of looking at things, particularly from a human rights lens and with an emphasis on the SDGs, which are valuable and should be maintained, including in support to NECD.

4.4.3 Evaluation and achievement of the SDGs

Agenda 2030 gave new impetus to NECD with its references to country-led evaluations, reinforced by the United Nations General Assembly Resolution. United Nations agencies are particularly well-positioned to strengthen the linkages between evaluation and progress towards the SDGs. Evaluation can inform improvements to strengthen the achievement of the SDGs by evaluating specific programmes or policies which contribute to the SDGs, evaluating sectors or whole SDGs, or using evaluative evidence in VNRs to inform the changes each country needs.

Arguably, Agenda 2030 has boosted NECD efforts with its references to country-led evaluations and building capacity for evaluation, reinforced by the United Nations General Assembly Resolution on evaluation capacity. Librado and Maclean, in their study of seven countries in Asia, highlight linkages between evaluation and the SDGs. They report a growing government commitment to evaluation in these countries, and awareness of its importance within the national development context. They all expressed strong political commitment to Agenda 2030 and had put in place institutional structures and mechanisms for mainstreaming the SDGs. Government leadership on SDG mainstreaming and evaluation was complemented by the commitment of many non-state actors, with increasing interest and demand from citizens for engagement both in evaluation and implementation, and tracking of country-level progress on the SDGs (Librado & Maclean, 2019).

From the perspective of UNEG members, the SDGs have helped to strengthen advocacy for evaluation. ILO shared that, by having a specific SDG (8) for which they were responsible, and then working this into national reporting, they had triggered requests for new support from their national constituents. Evaluation was promoted, as partners realized that it had a unique contribution to make, and that, while monitoring is important for reporting on *what* is happening, evaluation provides a voice on ***how*** things are working. Interestingly, UNICEF shared that, in one country, NECD actually provided them with an entry point to do work on the SDGs, where previously there had been no traction.

With respect to the Resolution, one United Nations respondent explained that, in the past, support for NECD in their agency depended on individuals, with a "known group" supporting it, while others were not interested. Even with the passing of the Resolution, there was not an immediate change, but after the Resolution was passed the headquarters office met with all of the regional advisers and participated in regional team meetings to raise awareness of the importance of NECD. With the passing of the

ILO training activities, Nepal 2016.

Resolution: *"it was no longer up to the individual team to decide."* NECD was now part of the mandate.

Yet views were not unanimous: some respondents to this study suggested that Agenda 2030 has not really boosted NECD efforts at country level 'as governments have seen that evaluation efforts are not necessarily linked to the SDGs.' Interestingly, one finding across the country case studies is that most respondents in country were not aware of the 2014 Resolution (69/237), and hence its biggest direct impact has been on United Nations agencies themselves, and only indirectly on countries.

UNEG member experience shows that there are several ways for evaluation to inform progress towards the SDGs. Key mechanisms by which evaluation can inform improvements to strengthen the achievement of the SDGs include: i) evaluation of specific programmes or policies which contribute to the SDGs; ii) evaluation of sectors or whole SDGs; or iii) using evaluative evidence in the VNRs to inform changes that the country needs.[32] With respect to the first, any evaluations which link to the specific SDGs and their indicators can make a contribution. In relation to the second, countries such as Benin are doing policy-level evaluations and some countries are undertaking evaluations of a whole SDG, for example, Nigeria has evaluated SDG 1 and is now evaluating SDG 5. However, the breadth of each SDG makes this difficult, especially if component programmes have not previously been evaluated. Uganda's evaluation map, mentioned previously, is organized by SDG and this enables a rapid use of

32 For example, following the 2020 VNR findings, the Federal Government of Nigeria has expanded the National Social Investment Programme to cover more poor and vulnerable people.

existing evaluation evidence to inform VNRs, or for sectoral meta-evaluations. United Nations agencies can support these SDG-oriented evaluations as part of their support to NECD.

With respect to VNRs, despite the Agenda 2030 call for review processes to be 'rigorous and based on evidence, informed by country-led evaluations' (United Nations: 2015), emphasis has been primarily on the use of statistics and on building national statistical capacity, with little use of evidence from evaluations in general. VNRs are mainly descriptive; they are not analytical and they do not help in understanding *why* things are (or are not) working and how situations can be improved. Using evaluative evidence offers significant opportunities to use VNRs in a more action-oriented way. Nigeria's 2020 VNR was one of the first to do this. It drew on 46 evaluations, using them to analyse blockages, and is action-oriented (OSSAP: 2020). UNICEF has been working with CLEAR-AA since 2018 to train African countries to embed evaluations in their VNRs, and this year has seen significant progress. For example, CLEAR worked with UNICEF and the Government of Madagascar to do a search and summary of evaluations, which has been an input to the country's 2021 VNR.

This challenge is mirrored in United Nations agencies themselves. Some agencies are custodians of reports on SDGs, and some evaluation offices report that they have struggled to collaborate with the department that prepares the report for the High-Level Political Forum on progress of the relevant SDG. Some suggest that UNEG could play a stronger role in this area.

4.5 Adapting evaluation to meet global crises

Support to NECD needs not only to strengthen traditional evaluation elements, but also to ensure that emerging national evaluation systems are prepared to meet the ever-growing challenges to the achievement of the SDGs. This requires information on what is (and is not) working to combat emerging crises, notably around climate and ecosystems emergencies, and the persistence of high levels of inequality.

Agenda 2030 clearly references the importance of protecting the planet from degradation and "taking urgent action on climate change, so that it can support the needs of the present and future generations" (United Nations, 2015). Thus, another question that arises is: to what extent are United Nations agencies taking the emerging global crises of climate and ecosystems breakdown and persistent inequality into account in their NECD work? These dimensions are critical for the survival of humanity, and many developing countries are very vulnerable to their impacts. Monitoring and evaluation have a role to play, and practices need to be transformed to respond to these complex issues. There could be a danger of supporting national evaluation systems which are designed to address historic, rather than future, challenges.

The COVID-19 pandemic has shown that major changes in society can be made if needed, and quickly. Some agencies reported that the climate crisis is now starting to

resonate: “While over the last two to three years, climate change started being taken more seriously, everyone still carried on travelling. COVID has changed that, and they have adapted so now thinking of the drastic changes needed around the climate crisis is more feasible.” Developing countries are likely to feel the impacts of climate change, ecosystems breakdown and inequality disproportionately, and this should be reflected in NECD work.

In terms of each agency’s own focus on **environmental sustainability**, some have conducted specific climate change-focused evaluations. For example, UNDP recently did an evaluation of its support to climate change adaptation, and OIOS will do a climate security evaluation in the next three years. However, there are more steps to be taken to update criteria to mainstream issues of equity, climate and ecosystems breakdown. IFAD has taken steps in this direction, integrating environment and natural resources management and adaptation to climate change into its criteria.

UNODC indicated that they are revising their Evaluation Policy to include a paragraph on mainstreaming environmental issues. WFP indicated that work is being done on mainstreaming climate change and many planned decentralized evaluations are touching on resilience and climate change as this has been prioritized by country directors. However, some agencies indicated that climate change, ecosystems breakdown and equity were not a priority in their evaluation work, even where these crises clearly affect the agency’s work. UNEG is taking steps to develop guidance on integrating environmental criteria into evaluation.

In terms of **inequality,** the integration of gender and disability considerations into United Nations work, including evaluation, is taken very seriously. There is a mandate and obligation to incorporate gender equality and disability into all evaluations, with agencies required to report annually against evaluation performance indicators.

These are important issues for evaluation, and by extension to ECD, if evaluation is to remain relevant. Reflecting on evaluation approaches and criteria is important as part of NECD discussions, given the absence of an equity criterion and the limited view of the sustainability criterion in the OECD Development Assistance Committee evaluation criteria, which are regarded in some evaluation quarters as very inadequate given the urgent crises facing humanity. This could result in evaluations looking at less significant issues while these crises gather pace.

Some agencies are providing support to integrate equity into SDG evaluations. For example, UN Women reported that they had provided technical assistance to Sri Lanka for evaluating the SDGs with and an equity-focused and gender-responsive lens.

The COVID-19 pandemic has forced innovation in evaluation methods, including the use of rapid methods. These methods need to be given more attention, as they are relevant to the emerging policy landscape characterized by volatility, uncertainty, complexity and ambiguity.

The COVID-19 pandemic has shown how evaluation methodologies can and do adapt to crisis, and how this can influence the methodologies promoted through NECD. Several agencies, such as OIOS, report that all of their evaluations have been conducted virtually for at least a year.[33] Some reported doing more rapid/ real-time evaluations. Others saw the increased use of big data, which is seen as a way to bring data in quickly. UN-Habitat raised the concern, however, that in times of crisis, funding for evaluation may be diverted, as evaluation is seen as too time-consuming and costly.

How are these innovations feeding through into NECD? The UNICEF/ CLEAR-AA training on embedding evaluation into VNRs has included training on rapid evaluations, evaluation syntheses and evaluative workshops. Twende Mbele has been supporting the training of African governments to undertake rapid evaluations, including Benin, Ghana, Kenya and South Africa.

In addition, the COVID-19 pandemic has pushed agencies to use more national capacity in their own evaluations, including the use of national evaluation teams rather than flying in international consultants, which can help to build the capacity of national evaluators. This could be part of a broader strategy to build national evaluation capacity. In general, agencies found this more cost-effective, with better contextualization of evaluations. Some indicated that it required more coaching, to make sure that data collection was going as planned. These lessons may make agencies more willing to try new approaches that develop capacity through the conduct of evaluations, as discussed above.

33 OIOS have developed evaluation guidelines for evaluating during times of crisis.

5. Conclusions and recommendations

5.1 Conclusions

UNEG has 53 diverse members, with mandates ranging from atomic energy to child poverty, from very big agencies with many evaluation staff to very small agencies where 'evaluation' is one person. Agency views on the UNEG role in NECD are thus also diverse, with some viewing NECD as very important and others, including those that arguably do not have the resources or country presence for significant investment in NECD, as a "non-priority."

However, Agenda 2030 and the decade for transformation make specific references to the importance of developing country capacity for evaluation. The 2014 General Assembly Resolution reaffirms the importance of national capacity for the evaluation of development activities. The UNEG Norms and Standards now include a Norm on national evaluation capacities.

Thus there is a call for the United Nations to invest across the board in NECD, and not just through small, individual agency efforts, but collectively, to advance human rights, governance, policies and budgets for those left behind. One United Nations respondent highlighted the importance of leadership and commitment: "the 2014 UN Resolution happened as there was strong leadership by a number of people who worked together in a concerted manner – without leadership NECD won't happen. This should be at all levels – director, technical. It's unlikely all 53 members will come together on NECD, perhaps this shouldn't even happen, but for those with capacity and a mandate, this should happen."

Evaluation capacity development should not be seen as an isolated set of activities. Rather, United Nations agencies should link NECD to broader evidence generation and use for planning and budget decision-making, which in turn will contribute to better development results. At the same time, stronger national evaluation systems and

Workshop with government officials during FAO country programme evaluation in Burundi, 2018

capacity will contribute to stronger and more useful agency evaluations. In this light, NECD can be seen as part of all agency mandates, which implies that a minimum part of every evaluation budget should be allocated to NECD. At the country level, and in line with the advances in United Nations reform which is all about working together, NECD should appear as a programmatic area within UNSDCFs, under governance, to bring agencies together.

With respect to progress at country level, evidence suggests that, in all of the case study countries, there has indeed been progress in NECD since the passage of the General Assembly Resolution in 2014. All countries have been supported to different degrees by UNICEF and UNDP, and a wide range of agencies have provided additional support. The case studies also show that building a national evaluation system is a long-term process. The multiple elements of the system need to be built progressively, and support needs to be consistent. Considering the importance of strong leadership and supporting champions, it is key to conduct evaluations even before all the elements of a national evaluation system are in place, as evaluations that are seen as useful can convince policymakers to support the process. This will also offer opportunities for government officials and non-governmental stakeholders to learn-by-doing. In addition, donor and United Nations funding for evaluations has the potential trigger government resourcing, as shown in Benin.

A lot can be achieved by adapting the way United Nations evaluations are conducted to enhance ownership and develop capacity through learning-by-doing. This may be, for example, ensuring that government representatives have meaningful roles in reference and steering groups, promoting joint and country-led evaluations with genuine country leadership and, when the systems exist, by using government terminology and systems. This will also ultimately strengthen the impact of evaluations

conducted by United Nations agencies. Applying a NECD lens to United Nations evaluations may require some adjustments. For example, making evaluation more responsive to partner country needs may mean adopting more flexible approaches, and/or focusing more effort on factors that support and encourage evaluation use. Harvesting lessons from successful uses of evaluation will also contribute to NECD.

There is room to create a shared framework for NECD work, and possibly to develop common work plans. There is not yet a shared language and understanding in UNEG of how to promote NECD, but more can be done even with limited resources. The 2012 UNEG Practical Tips Guidelines (UNEG, 2012) provide a good basis for this, and could be enhanced with the additional experience gained over the last ten years. There could also be a minimum standard among United Nations agencies to support country-led SDG-related evaluations.

UNEG could also provide resources on NECD, potentially in collaboration with other partners such as GEI. This could include: compiling and making public a repository of available resource materials/tools on NECD; distilling lessons about what works in NECD; and documenting and sharing case stories of successful NECD initiatives. These resources and tools could be shared with United Nations country team/country NECD focal points and be used for advocacy, conferences and discussions with partners outside of the evaluation community. In the words of one respondent:

> **"Ultimately, evaluation in the United Nations should include the vision we have for our partner countries: national evaluation being carried out of national programmes and feeding back into policy and practice related to the SDGs. These changes are possible."**

5.2 Recommendations

With a view to contributing to effective sustainable development and strengthening in-country evaluation capacity, United Nations agencies and their evaluation functions should implement the following recommendations, in line with their capacity and contexts.

Recommendation 1. All United Nations agencies should conduct their evaluations in a way that fosters national capacity development.

1.1 In principle, all United Nations agency country programme evaluations and UNSDCF evaluations should include the meaningful presence of national governments in management structures (reference groups, steering committees), with countries playing a leading role in such governance mechanisms. The next update of the UNSDCF evaluation guidelines should place stronger emphasis on supporting NECD and incorporate these recommendations into the text.

1.2 When feasible, United Nations agencies should foster joint and country-led evaluations. United Nations entities should also support country-led government

evaluations that address agency priorities, which may then count towards United Nations agency evaluation coverage requirements.

1.3 In countries with national evaluation systems, United Nations agencies should consider using national evaluation plans, guidelines, standards and other relevant elements in the conduct of their own evaluations, so as to respect the countries they are working in and enhance the credibility of these systems. In the process, this may help to further strengthen the national evaluation system.

1.4 United Nations agencies should commit to increase the numbers, and strengthen the capacity, of local evaluators, including through support to young emerging evaluators within evaluation teams. Agencies should consider using evaluation consultants as facilitators to build evaluation capacity, in addition to the standard evaluation responsibilities.

Recommendation 2. In line with General Assembly Resolution 69/237, United Nations agencies and their evaluation functions should continue to support the capacity development of national evaluation ecosystems, including support to the enabling environment, institutional and individual capacity. This may include a range of actions, aligned with the context and demand, as proposed in the sub-recommendations below:

2.1 United Nations agencies should, in collaboration with relevant partners and stakeholders, support country-owned M&E systems analysis to identify strengths and weaknesses in the broader ecosystem, followed by support to the definition and implementation of a medium-term evaluation capacity development strategy, according to their comparative advantages.

2.2 United Nations agencies should support the engagement of senior policymakers in the executive and parliament to increase their exposure to evidence-informed policy and practice, and respond to their evidence needs. United Nations agencies should support the development of a policy and regulatory environment to enable and sustain useful and credible evaluation processes and practices; as well as the strengthening of institutional capacity, frameworks and processes for conducting and using evaluations. Support for specific country-led or joint evaluations should be considered, to pilot-test instruments, promote opportunities for learning-by-doing, and demonstrate the usefulness of evaluation.

2.3 United Nations agencies should facilitate the engagement of non-state actors in the evaluation ecosystem, including VOPEs, academic and training institutions, citizens able to engage with evidence and policymaking debates, as well as evaluation professionals.

2.4 United Nations agencies should advocate for the integration of the SDGs, principles of gender equality, human rights, leave no one behind and disability inclusion, and climate change issues in country-led evaluations and national evaluation systems.

Recommendation 3. All United Nations agencies should coordinate and collaborate on NECD at corporate, regional and country levels, allocating adequate time and resources.

3.1 United Nations agencies should explicitly include NECD as part of their mandates, incorporate it into their evaluation policies, and allocate time and resources at corporate, regional and country levels. At least 10 percent of evaluation resources should be allocated to NECD.

3.2 United Nations agencies should ensure interagency information sharing, coordination and collaboration on NECD at corporate, regional and country levels.

3.3 At country level, United Nations agencies should include NECD as an explicit part of individual agency country programmes and of the UNSDCF, for example under a governance outcome, monitored by the United Nations country team M&E Working Group. When more than one agency is supporting NECD in the same country, they should coordinate efforts under a joint NECD programme, managed through a country-led steering group chaired by key government M&E champions and involving all actors in the evaluation ecosystem.

Bibliography

ADB. (1997). Technical Assistance to the Democratic Socialist Republic of Sri Lanka for Strengthening Performance Evaluation Capability of the Ministry of Plan Implementation, Ethnic Affairs, and National Integration. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/72389/30059-sri-tar.pdf

Amisi, M., Awal, M., Pabari, M., & Bedu-Addo, D. (2021). How relationship and dialogue facilitate evidence use. CLEAR-AA.

Association Marocaine de l'Évaluation. (2015). Memorandum sur l'institutionalization de la fonction de l'évaluation auprès des collectivités territoriales au Maroc. Rabat, Maroc.

Chaplowe, S., & Hejnowicz, A. (2021). Evaluating Outside the Box: Evaluation's Transformational Potential. Social Innovations Journal, 5. https://socialinnovationsjournal.com/index.php/sij/article/view/704

CLEAR Anglophone Africa. (2019). Embedding Evaluation in Voluntary National Reviews in Africa: A Guide. UNICEF/CLEAR-AA. https://evalsdgs.org/2020/01/01/guide-on-embedding-evaluation-in-vnrs-in-africa-enesfrpt/

CPRM Consultants. (2019). Review of National Evaluation Systems and Capacities for Evaluating Progress Towards the Sustainable Development Goals: Philippines Case Study. UNICEF/ UNDP.

Davids, M., Samuels, M. L., September, R., Moeng, T. L., Richter, L., Mabogoane, T. W., Goldman, I., & Buthelezi, T. (2015). The pilot evaluation for the National Evaluation System in South Africa – A diagnostic review of early childhood development. African Evaluation Journal, 3(1). https://doi.org/10.4102/aej.v3i1.141

Erasmus, Y., Jordaan, S., & Stewart, R. (2020). Scoping the impact evaluation capacity in sub-Saharan Africa. African Evaluation Journal, 8(1). https://doi.org/10.4102/aej.v8i1.473

Evalforward. (2020). Blog: Evaluation practice in Africa. Key trends in the landscape. https://www.evalforward.org/blog/evaluation-landscape-Africa

EvalForward. (2021). Blog: National ownership paramount - 5 lessons from a country-led evaluation in Peru. https://www.evalforward.org/blog/5-lessons-impact-evaluation-peru

Fakie, S, Lyby, E, Abedian, I, Vester, A, Kruger, F, du Plessis, P, Moran, G. (2009). 'Joint Evaluation of the Role and Contribution of the Unites Nations System in the Republic of South Africa, Pretoria, Government of South Africa/United Nations Evaluation Group.

Goldman, I., Byamugisha, A., Gounou, A., Smith, L. R., Ntakumba, S., Lubanga, T., Sossou, D., & Rot-Munstermann, K. (2018). The emergence of government evaluation

systems in Africa: The case of Benin, Uganda and South Africa. African Evaluation Journal, 6(1), 11. https://doi.org/10.4102/aej.v6i1.253

Goldman, I., Deliwe, C. N., Taylor, S., Ishmail, Z., Smith, L., Masangu, T., Adams, C., Wilson, G., Fraser, D., Griessel, A., Waller, C., Dumisa, S., Wyatt, A., & Robertsen, J. (2019). Evaluation2 – Evaluating the national evaluation system in South Africa: What has been achieved in the first 5 years? African Evaluation Journal, 7(1). https://doi.org/10.4102/aej.v7i1.400

Goldman, I., Olaleye, W., Ntakumba, S., Makgaba, M., & Waller, C. (2020). Mere compliance or learning – M&E culture in the public service of Benin, Uganda and South Africa. In Using Evidence in Policy and Practice—Lessons from Africa (First). Routledge, Taylor & Francis Group.

Goldman, I., & Pabari, M. (Eds.). (2020). Using evidence in policy and practice: Lessons from Africa. Routledge.

Government of Kenya. (2015). Concept Note: Formation of a Parliamentary Caucus on Evidence-Informed Oversight and Decision-Making.

IDEA International. (2018). Mapping Exercise on UNICEF Support to National Evaluation Capacity Development (NECD). UNICEF Evaluation Office.

IFAD Independent Office of Evaluation. (2015). Evaluation Manual second edition. IFAD Independent Office of Evaluation. https://www.ifad.org/documents/38714182/39748829/manual.pdf/bfec198c-62fd-46ff-abae-285d0e0709d6

ILO (2017). ILO Evaluation Policy.

ILO (2019). A brief history of evaluation in the ILO. http://www.ilo.ch/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_692334.pdf

Kenya National Treasury and Planning, State Department for Planning (2020). Monitoring and Evaluation Norms and Standards for the Public Sector.

Kenya Parliamentary Service Commission 11th Parliament. (2015). Concept note: Formation of a Parliamentary Caucus on Evidence-Informed Oversight and Decision-making.

Kenya State Department for Planning, Monitoring and Evaluation Department (2019) Guidelines for the Development of County Integrated Monitoring and Evaluation System. Council of Governors/The National Treasury and Planning. https://countytoolkit.devolution.go.ke/sites/default/files/resources/CIMES-Report_FINAL-28-Apr-1-1.pdf

Kenya State Department for Planning, Monitoring and Evaluation Directorate (2020) Kenya Evaluation Guidelines.

Kenya State Department for Planning, Monitoring and Evaluation Directorate (2021). Resolutions of the 9th Kenya Virtual M&E Week Conference 17th To 19th May, 2021.

Librado, D. & Maclean, M. (2019). Review of National Evaluation Systems and Capacities for Evaluating Progress Towards the Sustainable Development Goals in Asia and the Pacific: Synthesis Report of Indonesia, Malaysia, Mongolia, Philippines, Sri Lanka, Thailand, and Vietnam Case Studies. UNDP/UNICEF.

Manning, R., Goldman, I., Hernández Licona, G., & UNU-WIDER. (2020). The impact of impact evaluation: Are impact evaluation and impact evaluation synthesis contributing to evidence generation and use in low- and middle-income countries? (20th ed., Vol. 2020). UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/777-4

MIDEPLAN. (2020). Reporte de evaluaciones en el sistema nacional de Planificación 2019 – 2022: características. San José, Costa Rica. https://documentos.mideplan.go.cr/share/s/Cvbh4BMgTReO7KW7Ou-VBg

Morss, E. R. (1984). Institutional destruction resulting from donor and project proliferation in Sub-Saharan African countries. World Development, 12(4), 465–470. https://doi.org/10.1016/0305-750X(84)90024-X

Nigeria Office of the Senior Special Assistant to the President on SDGs. (2020). Nigeria: Integration of the SDGs into National Development Planning: A Second Voluntary National Review.

ONDH. (2015). Conférence internationale sur l'institutionalisation de l'évaluation des politiques publiques (Rapport de synthèse), Rabat, Maroc, 5-6 octobre 2015.

Picciotto, R. (2020). From disenchantment to renewal. Evaluation, 26(1), 49–60. https://doi.org/10.1177/1356389019897696

Primature/ Government of Benin, Direction Générale de l'évaluation. (2016). Rapport de suivi de l'utilisation des résultats des évaluations réalisées au cours de la période 2010–2013. Cotonou.
https://evaluation.gouv.bj/articles/28

South Africa DPME. (2020). Guidelines on Rapid Evaluation. https://evaluations.dpme.gov.za/images/gallery/GL%202.2.22%20Guideline%20on%20rapid%20Evaluative%20Processes.pdf

Todd, D. (2020). Stock-taking exercise on policies and guidance of United Nations agencies in support of evaluation of social and environmental considerations. UNEG Working Group on Integrating Environmental and Social Impact into Evaluations.

Trikawalagoda, P. (2018). Review of National Evaluation Systems and Capacities for Evaluating Progress towards the Sustainable Development Goals (p. 65). UNDP/ UNICEF.

UNDP. (2016). Country Programme Document for the Kingdom of Morocco (2017 – 2021), New York, US.

UNDP. (2019). Revised UNDP evaluation policy. http://web.undp.org/evaluation/documents/policy/2019/DP_2019_29_E.pdf

UNDP. (2021). Project Brief: Using Strategic Monitoring and Evaluation to Accelerate Implementation of the Philippine Development Plan 2017-2022.

UNDP IEO. (2019). Report of the Independent Evaluation Office on its support to evaluation capacity development. DP/2019/6.
https://undocs.org/DP/2019/6

UNDP IEO. (unpublished). UNDP Support to National Evaluation Capacity Development—Mapping Initial Findings. 2021.

UNEG. (2012). National Evaluation Capacity Development: Practical tips on how to strengthen National Evaluation Systems. https://www.unodc.org/documents/evaluation/Guidelines/National_Evaluation_Capacity_Development_Practical_Tips_1.pdf.

UNEG. (2016). UNEG Norms & Standards for Evaluation. http://www.unevaluation.org/document/download/2787

UNEG. (2019). UNEG Strategy 2020-2024.

UNEG. (2021). UNEG Partnership Strategy 2021.

UNHCR. (2016). UNHCR Policy on Evaluation. https://www.unhcr.org/research/eval/3d99a0f74/unhcrs-evaluation-policy.html

UNICEF Evaluation Office. (2018). Revised Evaluation Policy of UNICEF. https://www.unicef.org/evaluation/media/1411/file/Revised%20Policy%202018%20(Interactive).pdf

United Nations Economic and Social Council. (2020a). 2020 Secretary-General's Report on the implementation of the quadrennial comprehensive policy review. https://unsdg.un.org/sites/default/files/2020-06/SG-report-on-QCPR-implementation-30-April-2020.pdf

United Nations. (2011). Plan Cadre des Nations Unies pour l'Aide au Développement, UNDAF 2012 – 2016, Rabat, Maroc.

United Nations. (2015). Transforming our World: The 2030 Agenda for Sustainable Development.

United Nations. (2020). United Nations System-Wide Evaluation Policy - Shared Learning and Accountability for Delivering Results on Sustainable Development Goals. Draft 1.

United Nations Kenya. (2018). United Nations Development Assistance Framework. https://kenya.un.org/en/15986-undaf-2018-2022

UNODC. (2015). UNODC Evaluation Policy. https://www.unodc.org/documents/evaluation/Guidelines/UNODC_Evaluation_Policy.pdf

Waller, C. (2019). Baseline Study on the Performance Monitoring and Evaluation Culture in the Public Sector in Kenya (p. 35). Twende Mbele. https://twendembele.org/reports/baseline-study-on-the-performance-monitoring-and-evaluation-culture-in-the-public-sector-in-kenya/

WFP. (2016). WFP Evaluation Policy 2016-2021. https://documents.wfp.org/stellent/groups/public/documents/reports/wfp279331.pdf

WFP. (2021). EvaluVision: How visual thinking improves evaluation use and influence. https://www.wfp.org/publications/evaluvision-how-visual-thinking-improves-evaluation-use-and-influence

Appendix 1. Extracts from relevant United Nations resolutions

59/250 of 22 December 2004,

12. Recognizes that strengthening the role and capacity of the United Nations Development System to assist countries in achieving their development goals requires continuing improvement in its effectiveness, efficiency, coherence and impact, along with a significant increase in resources and an expansion of its resource base on a continuous, more predictable and assured basis. Section VII focuses on 'Evaluation of operational activities for development.

69. Encourages the United Nations Evaluation Group, under the aegis of the United Nations System Chief Executives Board for Coordination, to make further progress in system-wide collaboration on evaluation, in particular harmonization and simplification of methodologies, norms, standards and cycles of evaluation;

70. Strongly encourages country level evaluations of the (UNDAF) Framework at the end of the programming cycle, based on the results matrix of the Framework, with full participation and leadership of the recipient Government;

71. Recognizes that national Governments have primary responsibility for coordinating external assistance, including that from the United Nations system, and evaluating the impact of its contribution to national priorities;

62/208 of 19 December 2007,

10. Requests the United Nations Development System to continue its efforts to respond to national development plans, policies and priorities, which constitute the only viable frame of reference for programming operational activities at the country level, and to pursue full integration of operational activities for development at the country level with national planning and programming, under the leadership of national Governments, at all stages of the process, while ensuring the full involvement of all relevant stakeholders at the national level;

E. Evaluation of operational activities for development

129. Emphasizes that programme countries should have greater ownership and leadership in the evaluation of all forms of assistance, including that provided by the United Nations Development System, and requests the United Nations Development System to pursue and intensify its efforts to strengthen evaluation capacities in programme countries;

138. Encourages the United Nations Development System to further strengthen evaluation, with the agreement of the governing bodies of the funds, programmes and agencies, and in this regard encourages the United Nations Development System to continue efforts to strengthen evaluation across the system and to promote a culture of evaluation;

66/209 22 December 2011 (about supreme audit institutions)

67/226 of 21 December 2012

61. Calls upon the United Nations Development System to strengthen its focus on developing national capacities for development planning, disaggregated data collection and analysis, implementation, reporting, monitoring and evaluation, with an emphasis on the effective integration of the economic, environmental and social dimensions of sustainable development, and in this regard recognizes that the resources of the United Nations Development System, including the knowledge base and expertise of all resident and non-resident agencies, should be available for access by developing countries;

62. Also calls upon the United Nations Development System to further support the capacity-building and capacity development of developing countries, upon their request, and to effectively coordinate and evaluate the impact of external development assistance in line with national development plans and priorities;

F. Results-based management

164. Affirms the importance of results-based management as an essential element of accountability that can contribute to improved development outcomes and the achievement of the Millennium Development Goals and the internationally agreed development goals;

G. Evaluation of operational activities for development

175. Emphasizes that programme countries should have greater ownership and leadership of the evaluation of the assistance provided by the United Nations Development System, in this regard calls upon Members of the United Nations Development System to intensify efforts to assist programme countries to strengthen national evaluation capacity in programme countries for the monitoring and evaluation

of operational activities for development, and requests the United Nations Development System to develop and implement guidelines for further strengthening of national evaluation capacities for operational activities for development, in consultation with programme countries, including defining the responsibilities of different entities;

181. Requests the Secretary-General to establish an interim coordination mechanism for system-wide evaluation of operational activities for development of the United Nations system composed of the Joint Inspection Unit, the United Nations Evaluation Group, the Department of Economic and Social Affairs, the Office for the Coordination of Humanitarian Affairs and the Office of Internal Oversight Services, and also requests the Secretary-General, through the interim coordination mechanism, to develop a policy for independent system-wide evaluation of operational activities for development of the United Nations system, including submitting a proposal for pilot system-wide evaluations, for discussion by the Economic and Social Council at the operational activities segment of its substantive session in 2013;

Appendix 2. Survey of UNEG Members on National Evaluation Capacity Development

Background to the survey

The United Nations Evaluation Group (UNEG) National Evaluation Capacity Development (NECD) Working Group (WG) has been tasked with preparing a report on progress on implementation of United Nations General Assembly Resolution A/RES/69/237 on "Building capacity for the evaluation of development activities at the country level." This Resolution confirms the importance of building national capacity for the evaluation of development activities and invites United Nations agencies, with the collaboration of national and international stakeholders, to support efforts to further strengthen the capacity of Member States for evaluation in accordance with their national policies and priorities.

As part of undertaking this work we will be doing a survey of evaluation units in United Nations agencies to understand what you have been doing in the development of NECD. The questions relate to the theory of change for NECD being used by the UNEG Working Group. Please send your completed questionnaires to the consultant assisting us on this assignment, Professor Ian Goldman, at ian.goldman@wits.ac.za by 11 June. The closed questions should take you 20 minutes. The open questions are optional, if you don't have time.

Data privacy

The answers are indicative of the work of the organization rather than an individual and will be used as part of a database for the NECD Working Group.

1 Initiatives

We are attaching a document which shows the initiatives we already have for your agency. Are there any additional initiatives we should know about? Please list them below:

Initiative	Countries/region	Years	Weblinks

2 United Nations contributions to NECD (intended and unintended)

Please answer these questions for your agency as a whole.

2.1 Developing systems and capacity

2.1.1a	In the work you have been doing, was there a clear **demand for support** by government around evaluation?	**Yes**		**No**	
b	**Do you have any reflections on this?**				
2.1.2a	Have you supported building of **evaluation champions** in the centre of government and more widely?	**Yes**		**No**	
b	**In what way? Put an X in all that are relevant.**				
	Identifying individuals and giving them profile				
	Identifying individuals and exposing them to good practice in the country				
	Identifying individuals and exposing them to good practice outside the country through reports etc.				
	Identifying individuals and exposing them to good practice outside the country through visits				
	Providing training for those individuals				
	Other				

2.1.3a	Have you done work on **strengthening understanding in government of the role of evaluative evidence in policy and practice?**	**Yes**		**No**	
b	**If so, in what way?**				

2.1.4a	Have you done work on supporting the **drafting of evaluation or M&E policies and regulations**	**Yes**		**No**	
b	**If so, in what way?**				

2.1.5a	Have you done work on designing **evaluation plans, frameworks, systems and processes** to generate and use evaluative evidence?	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Evaluation Plans				
	Evaluation guidelines				
	Evaluation standards				
	Evaluation competences				
	Other (please specify)				

2.1.6a	Have you supported building **individual capacity to manage/conduct and use evaluative evidence**	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Training				
	Study tours				
	Learning-by-doing opportunities				
	Mentoring				
	Guidance				
	Internships				
	Other (please specify)				

2.1.7a	Have you supported **participation of non-government stakeholders** in the evaluation system	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Participation in structures overseeing evaluation in the country				
	Participation in evaluation steering committees				
	Participation in events guiding the TORs of evaluation				
	Participation in validation workshops on evaluation findings				
	Dialogues processes involving evaluations				
	Strengthening operation of the VOPE				
	Other (please specify)				

2.2 **Changes in outcomes** (behaviour and performance)

2.2.1a	Have you supported **production of evaluations by public institutions**	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Funding evaluations				
	Providing technical advice on conducting the evaluations				
	Participation in evaluation steering committees				
	Other (please specify)				
c	**How many have you supported from 2015?**				
d	**Did you make sure these would be publicly available?**	**Yes**		**No**	
e	**If so how? Put an x where appropriate**				
	Specifying in the funding agreement that these would be publicly available				
	Government putting them on their website				
	United Nations agency putting it on their website				
	Communication to the public about the availability of the evaluation				

2.2.2a	Have you implemented your **own evaluations in a way which builds the capacity of national systems**?	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Having these evaluations part of national evaluation plans				
	Joint funding with government of the evaluations				
	Joint funding with other agencies of the evaluations				
	Using government procurement systems to procure evaluators				
	Using government evaluation terminology and systems to undertake the evaluation				
	Using the evaluations to pilot with government how evaluation systems could be developed				
	Sharing the evaluations inside government or parliament				
	Other (please specify)				
c	**From 2015 how many evaluations have you implemented in a way which builds national capacity?**				

2.2.3a	Have you supported the production of **evaluations by non-public institutions**?	**Yes**		**No**	
b	**If so, in what way? Put an x where appropriate.**				
	Funding of the evaluations				
	Funding of processes to share evaluation findings				
	Other (please specify)				
2.2.3a	**Did you make sure these would be publicly available?**	**Yes**		**No**	
b	**If so how?**				

2.2.4a	Is there evidence that your agency has contributed to **increasing use of evaluative evidence in government decision-making?**	**Yes**		**No**	
b	**Would you say this is for (Put an x where appropriate)**				
	Most evaluations				
	Some evaluations				
	A few evaluations				
	Not happening very much				
	Other (please specify)				
c	**How have findings/recommendations been used? (put an x where appropriate)**				
	Findings/recommendations from government evaluations you supported have been implemented (i.e. used *instrumentally*)				
	Findings/recommendations from government evaluations you supported have built understanding of the problem, programme etc. (i.e. used *conceptually)*				
	The findings from government evaluations you supported have been used during the evaluation to strengthen the intervention or in other ways (*process* use)				
	The findings from civil society evaluations you supported have been used either instrumentally, conceptually or have had process uses				
	The evaluations have strengthened the interest in government in the use of evaluative evidence (*symbolic use)*				
d	**Do you monitor use in some systematic way?**	**Yes**		**No**	
	If so how?				

Are there any significant contributions your agency has made to NEWCD that you would like to flag?

3 Are there any overall lessons your agency has learned around NECD which you would like to flag?

4 What would you recommend going forward to strengthen how your agency supports NECD?

5 What would you recommend going forward to strengthen how UNEG supports NECD?

Thank you for taking the time to answer this survey. We hope it will contribute to improving evaluative practice in the United Nations< and how we work with our partner countries. Please send the completed survey to Ian Goldman at ian.goldman@wits.ac.za

Appendix 3. Members of Working Group

Name	Email	Organization	Interviewed
Renata Mirulla*	Renata.Mirulla@fao.org	FAO	X
Karen Bustamante	Ana.BustamanteAvendano@fao.org	FAO	
Aurelie Larmoyer	aurelie.larmoyer@fao.org	FAO	X
Patricia Vidal	vidalhurtado@ilo.org	ILO	X
Maite de Muller Barbat	mdemuller@iom.int	IOM	
Hanife Cakici	cakici@un.org	OIOS	X
Carlos Rodriguez Ariza	rodrigucar2@paho.org	PAHO	X
Michael Craft	Michael.Craft@unwomen.org	UN Women	X
Heather Bryant*	heather.bryant@undp.org	UNDP	X
Genta Konci	genta.konci@undp.org	UNDP	
Daniel Alonso	daniel.alonso@undp,org	UNDP	
Asela Kalugampitiya	kalugampitiya@unfpa.org	UNFPA	X
Oyuntsetseg Chuluundorj	oyuntsetseg@unfpa.org	UNFPA/ Asia Pacific Regional Office	X
Mahbub Alam	malam@unfpa.org	UNFPA/EECA Regional Office	X
Reginald Chima	chima@unfpa.org	UNFPA/ESARO Regional Office	
Martin Barugahare	Martin.Barugahare@un.org	UN-Habitat	X
Riccardo Polastro	rpolastro@unicef.org	UNICEF	X
Carlos Asenjo Ruiz	carlos.asenjoruiz@un.org	UNODC	X
Ivan Touza	ivan.touza@wfp.org	WFP	
Rica Terbeck-Soine	rica.terbeck@wfp.org	WFP	X
Michala Assankpon	michala.assankpon@wfp.org	WFP	X

*Co-chair

Appendix 4. Other people interviewed

Name (including title)	**Job-title**	**Organization**
Oscar Garcia	Director, Independent Evaluation Office	UNDP
Marco Segone	Director, Evaluation	UNFPA
Masahiro Igarashi	Director, Evaluation, OEDD	FAO
Andrea Cook	Director, Office of Evaluation	WFP
Simon Pierre Tegang	Evaluation adviser	Development Cooperation Office, United Nations
Ada Ocampo	President	IDEAS (and formerly UNICEF)
Dugan Fraser	Programme Manager	Global Evaluation Initiative
Tim Lubanga	M&E Commissioner	Office of the Prime Minister, Uganda
Angela Bester[34]	Consultant	

34 Interviewed as part of the evaluation team that undertook the evaluation of United Nations support to South Africa in 2008/9.

Benin interview respondents

Category	Name	Role and function	Organization
M&E champion	Prof Hygin Kakai	Vice-Dean Designed PPE Certificate course	Faculté de Droit et de Science Politique FADESP Faculty of Law Certificate in PPE
	Deo Gratias Houndolo	WACIE Coodinator	West Africa Capacity-building and Impact Evaluation Program
Line ministry M&E	Anicet Sevoh	Directeur Général Adjoint du Suivi des Investissements Publics	Ministre d'Etat chargé du Développement et de la Coordination de l'Action Gouvernementale
	Aristide Gnikpo	Directeur Adjoint de la Programmation et de la Prospective	Ministère de l'Agriculture de l'Élevage et de la Pêche
	Moise ILAYE	Ancien membre de la cellule de suivi-évaluation Ministère de la Décentralization Actuel Directeur de la Programmation et de la Prospective /Ministère de la Famille et des affaires sociales	Ministère de la Décentralization et de la Gouvernance Locale
Parliament	He. N'OUEMOU K. Domitien	Chairman, APNODE-Benin	Assemblée Nationale
	He. Adomahou Jérémie	Point Focal APNODE-Benin	Assemblée Nationale
United Nations agencies or other donors in the field working in NECD	José H. WABO	Deputy Resident Representative	UNDP
	Sylvano Nougbodé	Expert en appui au Suivi et à l'Évaluation du Programme Pays	UNDP
	Narcisse Kouthon	Chief PM&E UNICEF BENIN	UNICEF
	Koudoukpo Spéro	PM&E Specialist, PM&E Unit	UNICEF
	Cyrille Agossou	NPO, Chargé de Suivi & Évaluation	UNFPA
	Armelle Korogoné	Program Manager and Acting M&E manager	WFP
	Gérard Rubanda	Program Manager	WFP
	Aristide Djossou	Lead Economist, One UN	UNRCO Office
	Kamalou Moussa	Spécialiste Suivi-Évaluation	PAGIPG

Category	Name	Role and function	Organization
VOPE	Baudelaire Hounliho Secrétaire Général		Association des Professionnels de l'Analyse et de l'Évaluation des politiques
	Pamela AGBOZO	Secrétaire Générale Adjointe Cadre technique	Social Watch Bénin Réseau Béninois du Suivi et de l'Évaluation
Evaluator with experience of the system	Abdoulaye Gounou	Chef du Bureau	Bureau d'Évaluation des Politiques Publiques et de l'Analyse de l'Action Gouvernementale
	Damase Sossou	Evaluation Specialist	
	Elias Sègla	Evaluation Specialist	
Trainees Certificate in PPE (FGD)	Emmanuella Hekpazo		FADESP Faculty of Law Certificate in PPE
	Romaric Mouftaou		FADESP Faculty of Law Certificate in PPE

Morocco interview respondents

Name (including title)	**Job-title**	**Organization**
Mr. El Hasssan El Mansouri	General Secretary	ONDH
Ms. Chafika Affaq	Democratic Governance Programme Analyst	UNDP
Mr. Rachid AMRI	Planning, Monitoring and Evaluation Specialist	UNICEF
Mr. Carlos Asenjo	Evaluation Officer	UNODC
Mr. Ahmed Bencheikh	President of the Moroccan Association for Evaluation - AME	VOPE
Mr. Mohamed Abdouh	Professor and Director of the Centre for Regional Development (CEDRE)	Moulay Ismail University of Meknes
Mr. Mohamed Youbi Idrissi	Professor and President of the University Network of Inclusive Education - RUMI	Université Abdelmalek Essaâdi
Mr. Moustapha Boujrad	Consultant and co – founder of AME	VOPE
Mr. Ahmed CHAHBOUNI,	President	Centre de Développement de la Région de Tensift
Mr. Giovanni Saporiti	Consultant	Expert

Costa Rica interview respondents

Name	Job-title	Organization
Dña. Maria del Pilar Garrido Gonzalo	Minister	MIDEPLAN
Dña. Florita Azofeifa Monje	Director of the Monitoring and Evaluation Unit	MIDEPLAN
D. Eddy García Serrano	Chief of the Evaluation Unit	MIDEPLAN
M. Sc. Olman Villarreal	Director of the Post-Grade Program in Evaluation and Development Projects	UNIVERSITY OF COSTA RICA
Dña. Gabriela Pérez Yarahuan	Coordinator / Regional Director CLEAR LAC	CLEAR - CIDE
D. Juan Carlos Sanz	Project Evaluator, FOCELAC	DEVAL
Dña. Marta Villegas Murillo	Focal Point of Rural Development and Agroindustries Sector	FAO
D. Christian Vargas	Program Analyst P&D	UNFPA
Dña. Adriana Sánchez	Monitoring, Evaluation and Data Official	UNRCO
D. José Daniel Estrada	M&E Specialist	UNDP
Dña. Patricia Portela Souza	Country Representative	UNICEF
Dña. Michala Assankpon	Evaluation officer, Regional Office for Latin America and the Caribbean	WFP

Kenya interview respondents

Name (including title)	Job-title	Organization
Aloyce Ratemo	Director	M&E Directorate, State Department for Planning, The National Treasury and Planning,
Valerie Munyeti	RBM Analyst	UNDP Kenya
Rogers Dhliwayo	Economics Advisor	UNDP Kenya
Dan Juma	Team Leader Gov and Inclusive Growth Unit	UNDP Kenya
Mary Njoroge	Programme manager, devolution project	UNDP Kenya
James Wagala	M&E Specialist, Devolution	UNDP Kenya
Tim Colby	Technical Advisor, Devolution	UNDP Kenya
Obade Jackson Mukiri	Resource Mob Associate, doing a lot of M&E support	UNDP Kenya
Evelyn Koech	Team Leader, Environment and Resilience	UNDP Kenya
Jennifer Mutua	Past President	Evaluation Society of Kenya
Dr. Moses Ondabu Oyagi	Treasurer	Evaluation Society of Kenya
Julius Limbitu	CEO	Evaluation Society of Kenya
Kinlay Penjor	Research and Evaluation Specialist	UNICEF Kenya
Rhoda Goremucheche	Head of Capacity Development	CLEAR-AA

Sri Lanka interview respondents

Name (including title)	**Job-title**	**Organization**
Asela Kalugampitiya	Chair	Sri Lanka Evaluation Association
Ayanthi da Silva	Director General	Dept of Project Management and Monitoring
Sagarika Bogahawatta	Additional Director General	Dept of Project Management and Monitoring
Mohammed Ajiwadeen	Senior Research Officer	Parliament of Sri Lanka
Haana Singer	Resident Coordinator	United Nations Sri Lanka
Madura Chelliah	Data Management and Results Monitoring/Reporting officer	United Nations Sri Lanka
Rose Thompson Coon	Regional evaluation specialist covering & supporting Sri Lanka	UNICEF S Asia
Patricia Hurtado	Evaluation Officer	ILO Headquarters
Pamornrat Pringsulaka	Regional evaluation officer (co-convenor of UNIDAP)	ILO Regional Office

Philippines interview respondents

Name (including title)	**Job-title**	**Organization**
Xavier Foulquier	Chief PME	UNICEF Philippines
Lem Villamar	PM&E Officer	UNICEF Philippines
Koorosh Raffii	Regional Evaluation Advisor	UNICEF EAPRO
James Kimani	Multicountry evaluation specialist	UNICEF EAPRO
Ivan Scott	Consultant	UNICEF EAPRO
Marian Theresia Co,	RBM Analyst	UNDP Philippines
Maria Luisa (Lui) Jolongbayan	Institution and Partnerships Programme Team Leader	UNDP Philippines
Karen Dominique Brillantes	Institutions and Partnerships Programme Analyst	UNDP Philippines
Valerie (Val) Junginger	Strategic M&E Project Manager	UNDP Philippines
Kathkeen Ivy Custodio	Project Officer	UNDP Philippines
Romeo Santos	Evaluator	Member of PhilDev
Erika Lareza	Evaluator	Acting President, PhilDev

UNEG
United Nations Evaluation Group

http://www.unevaluation.org/