

The background features a large, abstract circular graphic composed of numerous overlapping, semi-transparent blue segments of varying shades, creating a dynamic, layered effect. The segments are arranged in a roughly circular pattern, with some extending towards the edges of the frame. The overall aesthetic is modern and digital.

Part III
Regional
Perspectives

Chapter 8

Evaluating Sustainable Development in SIDS - Lessons from the Pacific and the Caribbean

Juha I. Uitto, Jeremy Kohlitz, and David Todd

Abstract. *Small island developing states (SIDS) are particularly vulnerable to environmental stresses, and especially to the impacts of climate change. This is due to numerous factors, including limited geographic size and extensive coastal areas; remote locations; fragile economies that are often dependent on narrow sectors; limited natural resources and access to fresh water and energy; small populations; and weak institutional capacity. Managing sustainable development requires coherent and effective policies and strategies. An essential part of the formulation and implementation of such policies and strategies is effective monitoring and evaluation (M&E). Evaluation is also needed to ensure that interventions, policies, and strategies are achieving their goals and contributing to sustainable development. This chapter reviews experiences with M&E in the Pacific and Caribbean SIDS. It identifies a number of challenges that*

Juha I. Uitto, Global Environment Facility Independent Evaluation Office, juitto@thegef.org; Jeremy Kohlitz, Institute for Sustainable Futures, University of Technology Sydney, Jeremy.Kohlitz@uts.edu.au; David Todd, International Development, Environment and Disasters, deemtuk@yahoo.com.

need to be overcome, including limited human and institutional capacities, and the perception that evaluation is only important for donor-funded programs, which results in low priority being given to M&E. These challenges are best addressed by crafting M&E systems that are appropriate for a variety of SIDS contexts, that are country led, and that are supported by external agencies in a coherent manner.

Small island developing states (SIDS) are facing unique and often severe challenges to sustainable development. Their small size in terms of geography, economy, and population, and their limited capacities render them vulnerable to external shocks. The age-old limitations pertaining to natural resources, water, energy, and waste management have been exacerbated by global environmental change. Although their role in causing climate change has been minimal, they are at the frontlines of facing its impacts. Consequences ranging from increased weather variability and intensified storms to sea-level rise and salinization of groundwater pose serious threats to the sustainability of SIDS.

There are 57 countries classified by the United Nations (UN) as SIDS.¹ Most of them are located in the Caribbean and Pacific regions, but a number of them can also be found elsewhere—in the Atlantic and Indian Oceans, in the Mediterranean, and in the South China Sea. The UN has long recognized the special development situation of SIDS. The Barbados Programme of Action (BPOA), adopted in 1994 at the Global Conference on the Sustainable Development of Small Island Developing States, identified the unique nature of the vulnerabilities and characteristics of SIDS, including their small size, remoteness, and narrow resource and export base, as well as their exposure to global environmental challenges and to external economic shocks (UN 1994). Since then, a series of conferences under UN auspices has focused on devising concrete ways to further sustainable development in SIDS. For example, the 2005 Mauritius Strategy of Implementation (MSI) built upon the BPOA.² In 2014, the Third International Conference on SIDS was held in Apia, Samoa, and the SIDS Accelerated Modalities of Action (SAMOA) Pathways was adopted.³ Similarly, global processes have taken special note of the situation of SIDS. The 2012 UN report on “The Future We Want” dedicated a section to SIDS, noting that “small island developing States have made less progress than most other groupings, or even regressed, in economic terms especially in terms of poverty reduction and debt sustainability,” with member states reaffirming their commitment to providing assistance to implementing BPOA and MSI (UN 2012).

The 2030 Agenda for Sustainable Development, adopted by the UN member states in 2015, also focused on the specific plight of SIDS, especially in terms of climate change and associated issues related to sea-level rise,

¹ <https://sustainabledevelopment.un.org/topics/sids/list>.

² <https://sustainabledevelopment.un.org/conferences/msi2005>.

³ <http://www.sids2014.org/index.php?menu=1537>.

ocean acidification, and other impacts that are particularly affecting low-lying countries and coastal areas (UN 2015b). The attendant Sustainable Development Goals (SDGs), in particular Goal 13 (Climate Action) and Goal 14 (Life Below Water), are specifically relevant to SIDS. One of the challenges facing SIDS pertains to their generally limited capacities in terms of human and institutional resources. Consequently, the SDGs also call for “raising capacity for effective climate change-related planning and management.”⁴

Effective implementation of sustainable development strategies calls for effective monitoring and evaluation (M&E) in order to determine that processes are on track and that interventions, policies, and strategies are leading to desired change. Evaluating sustainable development in SIDS must take into account the economic, social, and environmental dimensions, while dealing with the considerable risk and uncertainties caused by global climate change, as well as possible discontinuities and tipping points in environmental trends.

Establishing effective M&E systems requires systematic effort and overcoming capacity constraints. An evaluation conducted by the Global Environment Facility Independent Evaluation Office (GEF IEO) in the Pacific found that all GEF projects have M&E protocols, and that the systems have been used effectively for adaptive management in the context of the projects. Yet institutionalizing M&E within the regular operations of the involved ministries and departments has proven challenging, primarily due to limited capacity (GEF IEO 2015). An evaluation in the Caribbean had similar conclusions: while project-level M&E has improved over time and has clearly contributed to adaptive management, environmental monitoring and the assessment of impact-level results have been extraordinarily challenging (GEF IEO 2012a). The reasons for this include a lack of baseline data as well as systematic monitoring data for assessing environmental trends over time. Other evaluations confirm these findings. For instance, in Timor-Leste and Jamaica, M&E has played a very limited role in managing the GEF portfolio and in providing environmental data to aid decision making (GEF IEO 2012b, 2012c).

As is evident from the analysis that follows, a number of factors hamper institutionalizing M&E in SIDS. One is simply the small size of governments and their capacities, combined with the fact that M&E is often seen as an external requirement related to donor-funded projects, thus rendering it low on the list of priorities. There are also significant challenges pertaining to data availability and the capacity to collect relevant data.

This chapter draws upon experiences in monitoring and evaluating sustainable development in the Pacific and Caribbean SIDS. It focuses on systemic and capacity constraints that need to be addressed in order to make M&E a useful tool for governments. In particular, evaluation is often seen mostly as being imposed by donor-funded projects or by regional organizations. It is important to overcome this perception in order to increase the utility of evaluation in these regions.

⁴Target 13.b.

SUSTAINABLE DEVELOPMENT IN SIDS IN THE ERA OF CLIMATE CHANGE

Recognition of the specific vulnerabilities of SIDS is not particularly new. In 1994, UNESCO's Island Agenda asked rhetorically, "Is Paradise an island?" (UNESCO 1994, 8), and outlined the various challenges facing small islands, ranging from small and subsistence economies to cultural issues and the mixed blessings of tourism. The report also identified limited natural resources, such as fresh water and energy, and the need for conservation of coastal and marine systems and unique island biodiversity as significant constraints to development. Vulnerability to natural disasters, such as cyclones, earthquakes, and tsunamis, and sea-level rise were identified as significant risks to small islands. Global warming, ocean circulation patterns, and climate variability were also mentioned, although at that time they did not yet receive major attention (UNESCO 1994). Similarly, the United Nations University World Institute for Development Economics Research presented a report to the Barbados Global Conference (UNU/WIDER 1995) that focused on overlapping issues. A chapter on natural disasters detailed the impacts on economic and social infrastructure in SIDS (Obasi 1995). Based on contemporary knowledge, the chapter was careful to note that there was no evidence that there had been an increase in tropical cyclones or their intensity due to climate change, but it also noted the increased risks associated with sea-level rise and its potential impacts on freshwater resources and coral reefs (Obasi 1995).

As the body of evidence grew, the international scientific community became more confident in stating that global climate change posed a major threat to low-lying coastal countries everywhere, and SIDS in particular. It was noted that SIDS were the first to pay the price for a problem that they had hardly contributed to (Pelling and Uitto 2001). Today we read reports of the dramatic effects of rising sea levels in island nations. An analysis of aerial and satellite images between 1947 and 2014 shows that in the archipelago of the Solomon Islands, five islands ranging in size from 1 to 5 hectares have already disappeared under rising seas, and another six islands have shrunk by 20–62 percent (Albert et al. 2016). Research appears to indicate that tropical cyclone frequency is decreasing, while the intensity of the storms appears to be increasing in ways that may create geomorphological change in the islands (Kelman 2016).

The 2015 United Nations report on SIDS focused exclusively on climate change, pulling together data from publicly available sources concerning the impacts on island nations (UN 2015a). The report identified serious threats to economic sectors that are essential for many islands, notably fisheries and marine resources, and tourism. It also highlighted how climate change impacts affect the social sector, including public health, food security, migration and displacement, and natural and cultural heritage. The report noted that the average annual losses from climate change are proportionally highest in SIDS: it is estimated that annual climate-related losses in Vanuatu are about 6.5 percent of gross domestic product (GDP). The cost of inaction is also high. If governments decide against any action toward climate change

adaptation, it is projected that the annual losses in the Caribbean will rise to \$22 billion—10 percent of the current size of the Caribbean economy—by 2050. In the Pacific, the total value of infrastructure, buildings, and cash crops at risk from climate change is estimated at \$111 billion (UN 2015a).

To address the above challenges, reliable data, systematic monitoring, and credible evaluation of the effectiveness and efficiency of policies, as well as their impacts, will be crucial. The following sections discuss issues pertaining to evaluation capacities in the Pacific and Caribbean island nations.

PACIFIC SIDS: DEVELOPING NATIONAL EVALUATION CAPACITY IN THE CONTEXT OF DIVERSITY

The nature of Pacific SIDS presents a special case for achieving and evaluating sustainable development.⁵ SIDS in the Pacific collectively comprise several hundred islands with remarkable geographic, socioeconomic, environmental, and cultural diversity, spread out over millions of square kilometers of the Pacific Ocean. Numerous threats throughout the region that have the potential to affect sustainable development include climate change; rapid urbanization that is spurring social inequality, deterioration of fragile biodiversity and ecosystem services, fresh water degradation, and resource depletion from growing populations (Duncan 2011; UNEP 2014). Furthermore, Pacific SIDS are often categorically described as being ill-equipped to face these challenges due to their remoteness, small size, and limited access to resources. However, the degree to which these and other factors affect sustainable development and the ability to react to change varies widely from island to island. When speaking of sustainable development in SIDS, it is important to be mindful that different SIDS face different challenges and opportunities. In the first of the following sections, the diversity of Pacific SIDS is emphasized, and the implications for evaluation policy are discussed.

Sustainable development requires sustainable evaluation: this is best achieved when countries have ownership and control of their own evaluation processes, as required by the SDG framework. The governments of Pacific SIDS face difficult challenges in owning evaluation: understanding their governance structures and processes is an important early step in planning how to fit an evaluation system into their specific contexts. The second of the following sections outlines the case for local evaluation of sustainable development, and provides a brief look at governance structures in Pacific SIDS.

Diversity of Pacific SIDS

A quick glance at geographic and demographic statistics of Pacific SIDS reveals stark differences between countries and territories. For example, the “small” in “small island developing states” has an indefinite meaning:

⁵ The Pacific SIDS are American Samoa, Cook Islands, Fiji, French Polynesia, Guam, Kiribati, Marshall Islands, Federated States of Micronesia, Nauru, New Caledonia, Niue, Northern Marianas, Palau, Papua New Guinea, Samoa, Solomon Islands, Timor-Leste, Tonga, Tuvalu, and Vanuatu.

land areas range from more than 450,000 square kilometers in Papua New Guinea to just 21 square kilometers in Nauru. The distribution of human settlements also varies widely, with the majority of Palauans and Marshallese living in central urban areas, whereas more than 70 percent of Ni-Vanuatu and Solomon Islanders live in rural areas spread across dozens of islands. People living on large volcanic islands often depend heavily on surface water sources and their associated ecosystems, while these do not exist on small atolls, where people depend instead on coastal and lagoon ecosystem services.

Levels of social and economic development also differ substantially. Five Pacific SIDS are classified by the UN as least developed countries—that is, countries that exhibit the lowest indicators of socioeconomic development. Meanwhile, some territories, such as New Caledonia and French Polynesia, are relatively developed, with a per capita GDP close to that of New Zealand.

Evaluation policy toward sustainable development in Pacific SIDS must be flexible enough to respond effectively to the variety of settings in which they exist. This requires representation from each of the Pacific SIDS so that they can explain their sustainable development objectives, and learn how evaluation resources can be made available to them, and used coherently. Fortunately, multiple sustainable development agendas for SIDS have already been put forward (e.g., BPOA, MSI, and the SAMOA Pathway), and attempts are being made to improve the platforms for regional cooperation in the Pacific, for example with the Pacific Plan and the Melanesian Spearhead Group (Pacific Plan Review 2013). These are useful entry points for the international evaluation community to promote and support the implementation of evaluation systems. These sustainable development agendas and regional platforms currently highlight the importance of evaluation, but there is room for more operationalization in the context of individual SIDS.

One evaluation challenge unique to some SIDS is the small size of governments. Pacific SIDS with very small population numbers often struggle to assess and report on all of the indicators set by external agencies. Human resources may be limited by the small number of government staff, or by a “brain drain” problem, as in the Cook Islands and Niue, whose inhabitants are automatically granted citizenship in developed countries that offer better economic opportunities. Careful thought must be given to prioritizing those SDG targets and indicators that can realistically be monitored.

Despite great diversity across Pacific SIDS, they do share important cultural characteristics that must be respected by external agencies that are looking to build their capacity for evaluation. Pacific SIDS have some of the highest levels of indigenous populations in the world, and among these populations, collectivist values are strongly maintained (Koshy, Mataki, and Lal 2011). As a result, many Pacific islanders strive toward achieving harmony, respect, risk avoidance, and loyalty within institutions even at the expense of efficiency and assertiveness as valued by Western institutions. There is a need to build capacity for evaluation in Pacific SIDS, but external facilitators must be sensitive to the way indigenous values affect the willingness and preferences of the islanders to develop this capacity. (For in-depth discussion of this topic, see Rhodes 2014).

Strengthening Government Capacity for Evaluation

Evaluation is more likely to be sustainable if the governments of SIDS are put in control of the evaluation processes and systems. Country- or government-led, as opposed to donor-led, evaluation is believed to help create a culture of evidence-based decision making, and to better reflect the information needs and values of the country stakeholders (Segone 2009). Establishing a sense of ownership over the systems can also help motivate government authorities to maintain evaluation processes, and to apply the results.

With respect to sustainable development, there are additional reasons to encourage government-led evaluation. First, not all sustainable development efforts are driven and funded by donors. For example, Pacific SIDS generally take the threat of climate change very seriously, and have their own sustainable adaptation initiatives and the associated information needs. Governments of Pacific SIDS should be supported in developing evaluation systems that they can use for their own endogenous purposes. Next, the social, economic, and environmental dimensions that make up sustainable development are adaptive and constantly changing. Unlike time-bound bilateral and multilateral programs, government is a permanent fixture, and it is in a better position to assess levels of sustainability over long periods of time if monitoring mechanisms are institutionalized.

Most sustainable development agendas influencing Pacific SIDS are conceptualized at the national or international levels, but governments will also need to think locally. Understanding interactions between human and environmental systems is key to assessing and evaluating sustainability (Rowe 2012), but these interactions are often highly context-specific. Therefore, in addition to top-down evaluation strategies, bottom-up strategies that can capture local nuances of sustainability are also needed. It will be beneficial to consider how local government in Pacific SIDS can be included and strengthened to fill this need.

Many Pacific SIDS governments have adopted decentralization policies, and several have constitutional provisions for local government. The scale of local government that exists varies widely (Hassall and Tipu 2008). In the Solomon Islands the local government is only decentralized to the level of provincial and municipal councils each of which currently oversee only tens of thousands of people. Meanwhile, individual islands and island groups in Kiribati have their own councils, some of which may oversee as little as a few hundred people. Some of the larger Melanesian SIDS have tiered national/regional/local government structures, while all government is maintained centrally in microstates like Niue and Nauru. Consideration of potential M&E responsibilities, and the mapping of the potential flow of information through disparate government structures is an important early step in assessing the potential for a comprehensive evaluation system.

Another interesting aspect of Pacific SIDS is that traditional or customary governance systems that date back to precolonial times are often blended with democratic governance systems and empowered through legislation (Hassall and Tipu 2008; Hassall et al. 2011). For example,

village councils composed of the heads of extended families are granted administrative power by the state in Samoa. In Tuvalu, elected local officials are accountable to a traditional assembly of elders who are given power by the state to oversee local affairs. These cases offer interesting opportunities for evaluation at the local level because the methods of inquiry could coalesce with traditional forms of engagement to produce rich, useful data that otherwise might be missed by top-down national approaches.

In order for evaluation systems to be effective within Pacific SIDS governments, they will need to be crafted appropriately to fit a variety of contexts, while working within on-the-ground realities. Sustainable development is a complex issue that plays out at not just at the global and national levels, but also at local levels, where traditional forms of governance are still common. Evaluation approaches need to be adaptive enough to handle both the complexity and variety found in Pacific SIDS, and should respect cultural heritage (Hoey 2015). Furthermore, local governments in Pacific SIDS generally are underresourced financially, and for the smallest islands this will likely continue for the foreseeable future (Hassall and Tipu 2008). The challenge going forward is not just to demand more evaluation and to offer training, but to work with Pacific SIDS stakeholders to identify which forms of evaluation make sense in each of their given contexts.

EVALUATING THE EFFECTS OF CLIMATE CHANGE ON SUSTAINABLE DEVELOPMENT IN CARIBBEAN SIDS

The SIDS of the Caribbean are among the world's most vulnerable countries when it comes to the effects of climate change (Todd 2011, 2013; UNFCCC 2005). This could become critical to their social and economic development, as well as to their terrestrial, coastal, and marine environments if no appropriate action is taken. Many islands are threatened by increases in the number and severity of extreme weather events, rising sea levels and coastal erosion, coral bleaching, and damage to biodiversity. At its worst, climate change could result in substantial loss of life and the damage to property and infrastructure that can easily cripple small and fragile economies. The Caribbean SIDS comprise a substantial part of the membership of the Caribbean Community (CARICOM), a regional political and economic grouping of some 20 member and associate member states. The World Bank has estimated that about 11 percent of the total GDP of all 20 CARICOM countries could be adversely impacted annually by climate change (Toba 2009; see also IDB 2014; World Bank 1997).

Although the Caribbean SIDS are not high greenhouse gas emitters on the world stage, they do have opportunities for climate change mitigation, and many countries are pursuing them through such activities as improved energy efficiency of buildings; increased production of solar power; the use of household equipment such as solar water heaters; and the scaled-up adoption of electric vehicles.

A Common Approach for Increasing Climate Change Resilience, Adaptation, and Mitigation among Caribbean SIDS

In 2009, through the Liliendaal Declaration,⁶ CARICOM leaders presented their vision of a common regional approach that would enhance resilience and adaptation by addressing the threats and challenges of climate change on Caribbean society and economy, as well as by providing support for mitigation-related policies and measures. This approach will be delivered through an overarching regional framework that comprises a set of strategic elements, each with defined goals and indicators, which should contribute to the achievement of the framework's overall objectives. Responsibility for designing and managing a common M&E system for the regional framework and its elements is vested in the Caribbean Community Centre for Climate Change (CCCCC), based in Belize.

The common approach is made up of five strategic elements that embody the key objectives contributing to the longer-term goal of building regional resilience and capacity to adapt to climate change:

- **Strategic Element 1.** *Mainstream climate change adaptation strategies into the sustainable development agendas of the CARICOM member states*
- **Strategic Element 2.** *Promote the implementation of specific adaptation measures to address key vulnerabilities in the region*
- **Strategic Element 3.** *Promote actions to reduce greenhouse gas emissions through fossil fuel reduction and conservation, and switching to renewable and cleaner energy sources*
- **Strategic Element 4.** *Promote actions to reduce the vulnerability of natural and human systems in CARICOM countries to the impacts of a changing climate*
- **Strategic Element 5.** *Promote actions to derive social, economic, and environmental benefits from the prudent management of standing forests in CARICOM countries*

This complex array of strategic elements has generated an associated set of 21 goals. It is intended that these should be monitored and evaluated in one M&E framework applied across the 20 member and associated states (Groupe-Conseil Baastel ltée 2013).

Challenges for Evaluation of the Regional Framework

Establishing, populating, and analyzing such an M&E framework would be a major undertaking for any region, and will pose particularly substantial challenges for many Caribbean SIDS, as outlined below.

⁶<https://caribbeanclimateblog.com/tag/liliendaal-declaration-on-climate-change-and-development/>

Challenge 1: Targets, indicators, and data. Progress toward each of the 21 goals should be measured through sets of baselines, targets, and indicators. The prevailing situation with respect to climate change targets is consistent across the CARICOM countries: that is, targets are generally only established for donor-funded projects. In many other cases, particularly for government-sponsored policies and activities, progress indicators have been defined, but are not associated with specific targets. Baseline data are generally available for national and sectoral-level frameworks, as well as for donor-funded projects, and are therefore relatively strong compared with targets and indicators.

Challenge 2: Variable resources for monitoring and evaluation. The data landscape is highly uneven across CARICOM member states. Countries with relatively major economies, notably Jamaica, have more resources to invest in national-level data collection and management than much smaller economies, where public administration has limited human and financial resources. Similarly, countries with many externally supported projects, such as St. Lucia, have greater access to external M&E advice, and the funds to support this function, than countries with fewer projects.

In recognition of this variation, it is important to note that M&E models that have been found effective in relatively well-resourced CARICOM countries cannot simply be transferred to other countries in the region, which have fewer resources. It is therefore important to ensure that evaluation activities and systems assess whether specific countries have made acceptable progress toward climate change targets according to their own unique situations, priorities, and resources, rather than against region-wide standards derived from countries with greater human and financial resources.

Challenge 3: Ability to respond to challenges of the regional framework. Underlying the concept of a unified CARICOM-wide M&E framework is the expectation that each country will have some indicators and targets that reflect the goals of its national priorities and policies. However, a preparatory study for the M&E framework found that many key stakeholders in government ministries dealing with climate change, or in other national bodies with a climate-related mandate, were not strongly aware of the specificities of the regional framework, and still less aware of how it might be evaluated. Furthermore, there is often no clear separation between monitoring activities and evaluation activities, which would only be feasible in much more developed and well-resourced systems. This means that even the limited data collected by monitoring systems are often not suitable for evaluation purposes.

Challenge 4: Low status of monitoring and evaluation. An additional important challenge to overcome is the low status and limited capacity of M&E activities in most of the countries in the region. Monitoring and evaluation are largely conducted in internationally funded interventions, and are often of a regional or subregional nature. Associated with such interventions,

several regional or subregional M&E frameworks have been proposed or established in various sectors (notably disaster management), usually with limited take up at the country level.

Underlying weak national implementation is the substantial inequality and perceived disconnect between the regional bodies proposing M&E frameworks and the national government departments or units that must conduct the detailed work of designing and implementing data collection and analysis at the country level. Characteristically, the regional bodies have sufficient technical capacity, human resources, and funding to participate in complex M&E exercises. Government departments, on the other hand, are often understaffed and poorly resourced, but have substantial implementation and reporting obligations, including some that are mandatory under the requirements of internationally funded activities or conventions. At the country level, therefore, evaluation is currently rarely recognized as useful, and has a correspondingly low functional status.

In the context of the constraints outlined above, the all-embracing M&E framework for climate change is often viewed as a largely unwelcome addition to existing tasks for the relevant government offices at the country level and may, in view of insufficient human and financial resources, be undeliverable. The requirements for national-level capacity building need to be carefully incorporated into the development of the overall regional M&E framework: this would include providing substantial financial resources over a period of time, given the low starting point in many countries.

Advancing Evaluation Capacity

The Caribbean SIDS have limited numbers of experienced evaluators, including those with climate change expertise. The demand for evaluators comes mainly from international bodies, including the countries and institutions that provide funding support to governments and regional institutions. To convert the desire for comprehensive evaluation of the effects of CARICOM policies and strategies to reality will therefore require substantial evaluation capacity development.

However, this capacity development needs to be placed within a fundamental upgrading of the role and implementation of evaluation. Underlying the possibility of such change are the following fundamental questions:

- How can the value and status of evaluation be raised, so that rather than being seen as a function that is of interest only to donors, it is seen as a useful resource for governments?
- How can evaluation support from various external funders be brought together into a coherent package, from its current state of fragmented bits and pieces?
- How can a more equitable and effective balance between the evaluation capacity of well-resourced regional institutions, and those of "shoestring" national government departments and offices be created?

If evaluation is to have any chance of becoming a viable component of the CARICOM regional approach to climate change adaptation, resilience, and mitigation, it needs to be very carefully focused on a few critical issues, and realistically scaled, so that it both appears to be, and is, implementable at the country level.

Evaluation capacity development needs to be part of a comprehensive package that will include in-person engagement of national stakeholders in its development through regional planning meetings of operational staff, and capacity building and financial support for national M&E functions. This process should place increased emphasis on evaluation, rather than focusing exclusively on indicators for results-based management.

FUTURE MONITORING AND EVALUATION IN THE SIDS

Experience in both the Pacific and Caribbean SIDS suggests that monitoring and evaluation as analytical tools are currently underutilized. However, we propose that M&E can be very powerful in promoting sustainable development in the island states, if it is institutionalized at the governmental level and integrated into government policies, strategies, and programs. Monitoring is needed to ensure that the various interventions are on track, and are completing their stated activities on time and in a cost-effective manner. Evaluation is needed in order to ascertain that the policies, strategies, and other interventions are reaching their goals and contributing to sustainable development without causing unanticipated negative consequences.

There are a number of prerequisites in order for this to happen. The first pertains to the fact that currently M&E is seen mostly as a requirement from donors and/or regional organizations. It consequently receives low priority at the level of national government ministries and departments. It is important to change this perception: this will require M&E functions and activities to demonstrate their added value. Building the capacity of national stakeholders to appraise and use evidence will contribute to creating demand for M&E. National stakeholders must be engaged in the development of these capacity-building efforts in order to promote country ownership, and a special focus on the utility of evaluation beyond monitoring indicators is needed.

There is also significant scope for cross-learning between SIDS in other regions: that is, not only in the Pacific and the Caribbean, but also in the Atlantic and Indian oceans, the Mediterranean, and the South China Sea. Existing mechanisms could be used for this purpose, including the Alliance of Small Island States (AOSIS), which has a membership of 44 states and observers covering all oceans and regions. Similarly, the UN manages the Inter-Agency Consultative Group (IACG) on SIDS, which brings together regional SIDS organizations⁷

⁷ Caribbean Community Secretariat (CARICOM), Commonwealth Secretariat, Global Island Partnership (GLISPA), Indian Ocean Commission (IOC), Organisation of Eastern Caribbean States (OECS), Pacific Islands Development Forum Secretariat, Pacific Islands Forum Secretariat (PIFS), Secretariat of the Pacific Community (SPC), and Secretariat of the Pacific Regional Environment Programme (SPREP).

as well as international, intergovernmental, and nongovernmental partners, and could ensure that all M&E activities pertaining to the agency work are coordinated. Another existing platform that could be used more effectively to share M&E experiences and lessons learned between the regions is the GEF International Waters Learning Exchange and Resource Network (IW:LEARN),⁸ which already provides an established forum for learning shared among a large number of environmental programs and projects across regions.

It is important to tailor M&E systems to the specific situations at hand. One size does not fit all, and top-down approaches that impose uniform frameworks on countries and departments regardless of their needs and capacities are counterproductive. It will be essential to choose the targets of monitoring and the subjects of evaluation carefully, based on the utility of these actions in helping to meet national priorities. It will also be important to adjust and design M&E systems to the particular institutional and cultural systems that are prevalent in each country.

Promoting country-led monitoring and evaluation will best serve the need for countries to manage sustainable development. Sustainable development issues involve balancing social, environmental, and economic costs and benefits: this means making value judgments about what exactly is most important to sustain in each country. Ultimately, it is the countries involved that should be making these judgments, informed by robust evidence and in recognition of the fact that sustainable development looks different at different scales and means different things to different people. Furthermore, climate change, which is one of the biggest threats to sustainable development in SIDS, is insidious because of the uncertainty about how society and nature will react to it in the coming decades. This makes consistent implementation and use of M&E all the more important: SIDS will need to be active learners as they adapt to constantly changing environments.

REFERENCES

- Albert, S., J.X. Leon, A.R. Grinham, J.A. Church, B.R. Gibbes, and C.D. Woodroffe. 2016. "Interactions between Sea-level Rise and Wave Exposure on Reef Island Dynamics in the Solomon Islands." *Environmental Research Letters* 11: 1–9.
- Duncan, D. 2011. *Freshwater Under Threat: Pacific Islands*. Bangkok: United Nations Environment Programme.
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2012a. *Cluster Country Portfolio Evaluation: GEF Beneficiary Countries of the OECS (1992–2011)*. Evaluation Report No. 72. Washington, DC: GEF IEO.
- . 2012b. *GEF Country Portfolio Study: Timor-Leste (2004–11)*. Evaluation Report No. 77. Washington, DC: GEF IEO.
- . 2012c. *GEF Country Portfolio Study: Jamaica (1994–2010)*. Evaluation Report No. 66. Washington, DC: GEF IEO.
- . 2015. *GEF Country Portfolio Evaluation: Vanuatu and SPREP (1991–2012)*. Evaluation Report No. 98. Washington, DC: GEF IEO.

⁸<http://www.iwlearn.net>.

- Groupe-Conseil Baastel Itée. 2013. "Monitoring and Evaluation Instrument and Reporting Framework in the Caribbean." Prepared for the Caribbean Community Centre for Climate Change.
- Hassall, G., and F. Tipu. 2008. "Local Government in the South Pacific Islands." *Commonwealth Journal of Local Governance* 1: 6–29.
- Hassall, G., M. Kaitani, F. Tipu, and L. Wainiskesa. 2011. *Comparative Study on Local Government and Traditional Governance Systems Interaction in Pacific Island Countries*. Suva: University of the South Pacific.
- Hoey, L. 2015. "Show Me the Numbers: Examining the Dynamics between Evaluation and Government Performance in Developing Countries." *World Development* 70: 1–12.
- IDB (Inter-American Development Bank). 2014. *Building Resilience and Reducing Emissions*. Office of Evaluation and Oversight. Washington, DC: IDB.
- Kelman, I. 2016. "Science Shows Climate Change Impact in the Pacific." *Australian Journal of Emergency Management* 32 (1): 15–16.
- Koshy, K., M. Mataki, and M. Lal. 2011. *Sustainable Development—A Pacific Islands Perspective*. Apia: UNESCO.
- Obasi, G.O.P. 1995. "Natural Disasters and Sustainable Development of Small Developing Islands." In *Small Islands, Big Issues: Crucial Issues in the Sustainable Development of Small Developing Islands*, United Nations University World Institute for Development Economics Research (UN/WIDER), 63–83. World Development Studies 1. Helsinki: UN/WIDER.
- Pacific Plan Review. 2013. "Report to Pacific Leaders." Pacific Islands Forum Secretariat, Suva.
- Pelling, M., and J.I. Uitto. 2001. "Small Island Developing States: Natural Disaster Vulnerability and Global Change." *Environmental Hazards: Human and Policy Dimensions* 3 (2): 49–62.
- Rhodes, D. 2014. *Capacity across Cultures: Global Lessons from Pacific Experiences*. Fairfield: Inkshed.
- Rowe, A. 2012. "Evaluation of Natural Resource Interventions." *American Journal of Evaluation* 33 (3): 382–92.
- Segone, Marco. 2009. "Enhancing Evidence-based Policy Making through Country-led M&E Systems." In *Country-led Monitoring and Evaluation Systems*, Marco Segone, ed. New York: UNICEF.
- Toba, N. 2009. "Potential Economic Impacts of Climate Change in the Caribbean Community." LCR Sustainable Development Working Paper No. 32. World Bank, Washington, DC.
- Todd, H. 2011. "Children and Climate Change: Action for Prevention, Mitigation and Response to Fulfil the Rights of Children to Education in Emergencies." Final report submitted to UNICEF, Barbados.
- . 2013. "Disaster Risk Management, Climate Change and Food Security in Barbados." Final Report to the Food and Agriculture Organization of the United Nations.
- UN (United Nations). 1994. "Report of the Global Conference on the Sustainable Development of Small Island Developing States." A/CONF.167/9.
- . 2012. "The Future We Want." A/RES/66/288.
- . 2015a. *Small Island Developing States in Numbers: Climate Change Edition*. New York: Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States (UN-OHRLLS).
- . 2015b. "Transforming Our World: The 2030 Agenda for Sustainable Development." A/RES/70/1. UN General Assembly, New York.

- UNEP (United Nations Environment Programme). 2014. *Emerging Issues for Small Island Developing States*. Nairobi: UNEP.
- UNESCO (United Nations Educational, Scientific and Cultural Organization). 1994. *Island Agenda: An Overview of UNESCO's Work on Island Environments, Territories and Societies*. Paris: UNESCO.
- UNFCCC (United Nations Framework Convention on Climate Change). 2005. *Climate Change, Small Island Developing States*. Bonn: UNFCCC Secretariat.
- UNU/WIDER (United Nations University World Institute for Development Economics Research). 1995. *Small Islands, Big Issues: Crucial Issues in the Sustainable Development of Small Developing Islands*. World Development Studies 1. Helsinki: UNU/WIDER.
- World Bank. 1997. "Caribbean Planning for Adaptation to Global Climate Change Project." Project document. World Bank, Washington, DC.

Chapter 9

Evaluation Standards for Latin America and the Caribbean - Experimentation and Evolution

Ana Luisa Guzmán and Warren Crowther

Abstract. *This chapter reviews an excellent contribution to the development of evaluation standards for Latin America and the Caribbean developed by the Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC), the Evaluation Capacity Development in Selected Countries in Latin America (FOCEVAL), the German Institute for Development Evaluation (DeVAL), and the Costa Rica Ministry of National Planning and Political Economics (MIDEPLAN). Lessons learned from experimentation in applying a code of ethics in the development of seven diverse and ambitious projects over a 10-year period in Costa Rica are discussed. The basic principles covered by the standards are posited, and the practical relevance, viability, and trade-offs, or relative emphasis in the evaluation and decision-making aspects of the projects, are analyzed. The proposed standards for Latin America and the Caribbean constitute an excellent guideline for applying and implementing the improvement of evaluation processes. This observation and analysis suggests an agenda for further debate, including the inclusion of additional evaluation standards.*

Ana Luisa Guzmán, Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC) and University of Costa Rica, guzmaneste@hotmail.com; Warren Crowther, Independent Consultant, warren13@racsa.co.cr.

The Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC), the Evaluation Capacity Development in Selected Countries in Latin America (FOCEVAL), the German Institute for Development Evaluation (DeVAL), and the Costa Rica Ministry of National Planning and Political Economics (MIDEPLAN) have all made an important contribution to the experimentation and evolution of standards for program evaluation in the region. In the first stage, many evaluators were consulted and a literature review was carried out regarding competencies and standards, including the evaluation standards used in United Nations (UN) agencies, internationally recognized partnerships, and other such initiatives. In 2015, a draft proposal of standards was circulated among and commented upon by interested members of the network, and in 2016 a published edition of the Evaluation Standards for Latin America and the Caribbean (EEALC; Rodríguez et al. 2016) was circulated and commented on by interested organizations and professionals.

During this period of elaborating and sharing criteria for the standards, a review of elements of ethical guidance—based on the code of ethics published by UNICEF (2002) for research, monitoring, and evaluation of programs and projects taking into account the rights and interests of children and adolescents—took place in Costa Rica.¹ Its basic principles can be applied to the target populations in the design and evaluation of the programs and projects catering to their rights and interests, and the use of action research methodologies. These same principles were adapted for other target populations, including low-income families needing assistance with initiatives in entrepreneurship, exceptional patients with serious illnesses, older adults, and communities and users affected by transportation and communications investments.

The experience with these projects is instructive regarding the need and conditions for an adequate set of standards, adapted to Latin American and Caribbean realities, while recognizing that although Costa Rica has a mixed population when it comes to ethnicity, urban and rural culture, economic conditions, and environment, it still represents only a fraction of the diversity that needs to be considered in the entire region.

The lessons learned from the demonstration projects are summarized in table 9.1. They highlight implications for the evolution of program evaluation standards for the region.

The most interesting of these experiences illustrate the ample options and major obstacles to some of the critical factors of quality monitoring and evaluation (M&E) of policies, programs, and projects (PPP), which not circumstantially have received the most intensive attention and development in the discussions regarding the EEALC.

Between the initial proposal and the latest version of the EEALC, there has been an impressive and promising advancement in the definition of these standards, particularly related to four major themes: relevance to decision

¹ In addition to UNICEF (2002), attention was also given to two other norms: UNICEF Costa Rica (2000) and ICAP and UNDP (2003).

making; influence on decision making; validity of the evidence; and the fulfillment of preconditions. Table 9.1 illustrates the importance and value of these themes, as well as some nuances or complications that have not yet been broached in the standards, and are in constant evolution as their application is being monitored.

Some of the projects are reviewed in table 9.1, regarding what was intended, what resulted, and what was learned. These are all demonstration projects, initiated by the ProHumana21 Foundation between 2004 and 2014. They were undertaken with intensive grassroots participation and attempts to create prototypes or models of better practices, including more adequate impact assessment.

USING EVALUATION FOR DECISION MAKING

This is mentioned in the EEALC: it is an important starting point, but it needs to be reviewed in order to consider programs and projects as subject to constant experimentation, and not just implementation and resolution of pre-determined terms of reference and premises. There are always both positive and negative surprises that affect the engagement, dimensions, and parameters of the M&E and decision-making processes of projects and their impacts. The projects reviewed in table 9.1 emphasize this.

It became apparent that in the implementation of these projects, there was no relationship between the relative gravity or threat of the situation discovered with innovative evaluation methods, and the amount of local and/or national interest that existed to deal with those facets of that situation, and in defining the projects' terms of reference. Thus a focus on the unexpected turned out to be more revealing—in either an instructive or a disturbing way—than a focus on the achievement of pre-established goals.

For example, it was observed that the resolution of many socio-environmental conflicts and medical pathologies led to unanticipated delayed effects in the form of new conflicts or pathologies. The apparent resolution of one socio-environmental conflict often leads to the unraveling of other conflicts, and the remission of one illness is often the preface to a later downturn in health. Thus, the adoption of a solution for one challenge can often create another challenge.

Another insight from analyzing these projects resulted in fostering comparative research and evaluation, which involved ambiguity in choosing the models and indicators that were sensitive to particular cultural and local realities, and those that were emphasized when comparisons were made. Yet, ironically, comparisons can help rather than hurt in this situation. The greatest "relevance" of an evaluation is not only in appraising how well preconceived challenges and the consequent predefined goals are being met, but also in identifying or uncovering the locally or presently unseen or unexpected factors: in this regard, comparative evaluation can actually be very helpful in making those factors more visible or well anticipated.

The periodic Latin American Development Administration Congresses (CLAD) are very instructive in comparing experiences: however, in the most recent congress, held in Chile in 2016, there was a lack of communication

between two main streams of work. There were several sessions dedicated to reviewing advances in PPP evaluation, including discussion of the EEALC. Other sessions were focused on program and project “innovation and government laboratory” advancements: yet it was noticeable that very few people who attended the evaluation sessions also attended the innovation sessions. Unfortunately, there was no discussion about the importance of assimilating the benefits of these approaches. We will readdress this perspective when we consider the importance of evaluating impact and not just results.

HOW EVALUATION AFFECTS DECISION MAKING

The projects in Costa Rica have used action research methodologies to integrate, at the local level, the evaluation and decision-making processes. These projects have also invited—or pressured and educated—the target populations and other present stakeholders who sometimes met with resistance, to broaden their perspective on the relationship between evaluation and decision making, and how to most positively engage in them. Attention was given to questions of confidentiality and sensitivity of the evidence, and judgments in the evaluation, which prior to the evaluation should be agreed upon by key stakeholders.

In the projects reviewed in table 9.1, a general benefit was enhancement of the knowledge management capacities of the most active among the target populations, and at times of other stakeholders as well. However, this meant focusing more attention on the sensitivities—and sometimes uncomfortable ethical implications—of this knowledge management sharing and decision-making responsibility and privilege, and of the evidence itself. A major difficulty arises where a “culture of shame” prevails, meaning a general avoidance of sharing what could be sensitive or embarrassing information.

The common withholding or suppression of evaluation results was even more tense and awkward: this included evidence considered by one or more parties to be offensive to their agenda or interests, or potentially damaging to other parties.

LIMITATIONS OF EVIDENCE VALIDITY

In the projects reviewed in table 9.1, it became evident that the more intimate and closely connected the evaluators were to the local details and consequences of PPP implementation, the greater the difficulties they faced in identifying evaluation indicators that were both suggestive of validity and of applicable practicality in the evaluation. This was also illustrated by empirical results highlighted throughout the analyzed literature review. That is, the indicators were more geared to their relatively easier application and to acceptability by data sources than to the pertinence of the evidence for the variables that were posited in each case. This review depicted inconsistencies between the connotations of the variables outlined in the conceptual or modeling configuration of what was being studied or evaluated, and what was evidenced through the practical indicators.

Pertinent to this are proposals regarding quality criteria for the selection of indicators. Somewhat different examples of such proposals have been offered by Stockmann and Crowther, both of whom have shown that “ample criteria” included some that in practice are not compatible with each other (Crowther 1999; Stockmann 2011).

This means that inevitably, in all evaluations, the predispositions of the evaluator (or decision makers) play a role regarding the overall criteria used for selecting indicators.

THE REQUIRED PRECONDITIONS FOR A PARTICIPATORY EVALUATION

For almost any reputable and useful evaluation, considerable work has to be done just to create the preconditions. This has been the case for decades in Latin America, where the insistence on a precipitated or institutionalized evaluation procedure without these preconditions has come to denigrate the credibility of evaluation. Often those most affected—that is, either those being most closely evaluated or the major sources of evidence—might see program evaluation as, for example, lying, “tattling,” an intrusion. Or they might simply resist it altogether.

OTHER CONSIDERATIONS AND ISSUES

Different initiatives for developing evaluation standards have taken various positions regarding the four major themes, and constructive attention has been given to them in EEALC.

It is notable that with the exception of the requisite conditions for the evaluation, the United Nations Evaluation Group (UNEG), in its 13 norms, tends to assign less attention to these four standards, and instead highlights the ones that EEALC has identified as the most difficult ones to resolve. For example, UNEG stresses impartiality, and many in UNEG would see impartiality as supporting and strengthening validity. However, UNEG is not explicit about the inherent limitations of evidence and judgment validity for example, regarding issues of sensitivity and confidentiality. On this last point, the American Evaluation Association’s Program Evaluation Standards more realistically assigns importance to “technically adequate” information.

In applying these standards, a general lesson from the projects shown in table 9.1 is that their compliance is fostered with the application of action research principles and methodology. In addition to complex planning and implementation of the evaluation process, more than the usual amount of human and economic resource dedication is also required. The target populations and other stakeholders are not readily prepared for the protagonist role they are invited to, or that they need to, assume: much preparation for this role is needed.

On the other hand, there are five more themes with which the projects faced more troublesome practical challenges, and for which, optimistically, the evolution of EEALC may provide solutions. These themes are: involvement of the principal stakeholders in decisions about the evaluations; examination

of the underlying values influencing the evaluations; impact; evaluator competencies; and the role of context.

DECISION MAKING: WHO DECIDES WHAT?

The EEALC indicates that the main target population of the PPP are “groups affected by” the evaluation. Regarding the very diverse sets of responsibilities and/or privileges they should have in the evaluation and decision-making process, 16 different categories or segments of actors or stakeholders are designated, with some overlapping regarding both groupings and functions.

For the projects shown in table 9.1, the most interested segments of the population could be readily identified. However, some of the interested stakeholders demonstrated some degree of resistance to sharing or conceding their personal incidence in the decision making or real, informed participation in the design and evaluation of PPP, to those most affected by them. This resistance was less than expected, although it varied among localities and subcultures. With PPPs directed at the rights and interests of adolescents and youth, there was less adult centralism than anticipated, but there was much more than expected regarding PPPs directed at older adults, including among the professionals attending this generation.

UNDERLYING VALUES, OR CRITERIA, INFLUENCE WHAT IS BEING EVALUATED

In the projects reviewed in table 9.1, the agendas of key stakeholders often did not coincide with the ethical standards the projects were trying to adopt: or they were not compatible with each other regarding the underlying values or criteria to be applied.

In these projects, there were numerous cases of unpreparedness for new challenges, and others of institutional units adhering to their agendas and general priorities, and seeking support for more coverage and technological upgrading rather than pertinence and realism in what they were offering. In such cases, substantive evaluation could seem very threatening.

The inherent or inevitable incidence of the evaluator’s (or decision maker’s) predisposition regarding the selection of indicators is presented in all areas of evaluations, since an external referent to the situation being evaluated is always being introduced either implicitly or explicitly -reporting that things go well or badly, better or worse, optimistically or fearfully, and so on. (Perhaps the exception to this is the response “accepted or tolerated” or “not accepted or tolerated.”) One can adopt a baseline or benchmark to show apparent progress or regression, but there is always that important question, which is usually unanswered—“In comparison to what?”—which is needed in order to conclude whether a “reasonable expectation” is being met. Where does this “reasonable expectation” come from? Not only is the lack of reference to a defined reasonable expectation a common deficit in evaluation: it is also lacking in almost all social, educational, and administrative research.

TOWARD EVALUATING IMPACT

When comparing proposed evaluation standards, it is important to consider the connotations and meanings in different languages of the terminology being adopted. In fact, the same term can have different connotations even in different countries that speak the same language. That said, in the evolution of regional standards in Latin America and the Caribbean, the most consistently affirmed values, or criteria, have been utility, precision, pertinence, and timeliness.

A most concerning tendency in the region is placing much more emphasis on evaluating *results* than on evaluating *impact*, although increasing attention is being given to the latter. The positive side of the emphasis on results is that it gives more attention to the democratically elected government's compliance to the promises they have posited, or the expectations they have raised. But a downside is that it can lead to efforts to improve a PPP, or do a better job of managing it, when a real impact evaluation may indicate that it shouldn't be taking place at all, or that it needs major reform.

Impact as such is not mentioned in the EEALC, but there are interesting hints of it in the insistence that there should be a positive contribution to decision making, and the "quality of life" of the "beneficial public" and "other interested parties."

The UNEG norms make seven references to the evaluation of impact as an alternative to evaluating outcomes: output, relevance, efficiency, effectiveness, sustainability, value-for-money, and client satisfaction. But those norms don't really consider what impact as such entails.

For the projects summarized in table 9.1, a strong definition of impact has been adopted, which includes the added value of the target population regarding the most fundamental transformations of its quality of life; the target population having good knowledge of and insight into what *constitutes* and influences the quality of their lives; and consideration of the opportunity costs of all interested or affected stakeholders. This definition is not very operative, except as a point of reference to discuss with the key stakeholders, including the target populations. It means not just comparing yesterday or today with tomorrow, which is the usual manifestation adopted for a baseline: this kind of benchmarking needs to project *what was, or is, likely to occur* in a given period of time *without* the intervention that is being evaluated.

WHAT COMPETENCIES SHOULD THE EVALUATOR HAVE?

This is a delicate topic: many people who have engaged in evaluation diligently and continuously do not begin to have the qualifications elaborated on in texts such as that of Rodriguez et al. (2016). Furthermore, some of the standard qualifications are extremely difficult to evaluate, such as whether an evaluator is being "objective" and "impartial." The emphasis on credibility puts this up for grabs: who will be most convincing to those who will be hiring the services of an evaluator? Comparing the terms of reference of such

qualifications—including for other projects with the same target populations as those in table 9.1—the differences as to expectations of competencies are monumental.

TAKING CONTEXT INTO ACCOUNT

The importance assigned to “context” is often associated with recognition and respect for local and cultural differences. What is seldom discussed is how much subjectivity is inevitable in determining what “context” to consider, as almost all of the phenomena being evaluated are interrelated with infinite conditions or variables, and there is always a presumption as to where and how to draw the lines.

For Costa Rica and the other 40 countries categorized as high middle income, there is an additional limitation: not counting on think tanks dedicated to analyzing and comparing their commonalities, such as those that exist for the more developed and the less developed countries. There is little systematization of the common problems at this stage of development, which could help contextualize studies in the individual countries.

CONCLUSIONS

It is extremely challenging to create a consensus about more precise standards especially regarding the last five mentioned above—decision making, the influence of underlying values or criteria, evaluating impact, evaluator competencies, and context—including an operative definition of impact.

Nonetheless, UNEG’s 13 norms and the 30 standards of the American Evaluation Association’s Program Evaluation Standards focus on these very expectations, in some cases mentioning them without really defining them.

These international norms and standards also give major attention to two issues that can be closely interrelated, and which the Latin American and Caribbean standards do not deal with directly: that is, human rights in general, and the well-being of those participating in the evaluation, not just those affected by its results. Human rights, and especially its byproducts of equity or equality, involve constant and inherent trade-offs. Fourteen operative definitions of “equity” that are commonly adopted in practice can be identified; all have substantial legitimacy, but they can also be quite contradictory with each other.

It is necessary to recognize the importance of the learning experiences presented in this work, both in terms of the results and the impacts obtained in the projects developed, and the constant search for high quality in their evaluation. This is especially true in the case of Latin America, where it will help to refine the evaluation standards with greater precision regarding the contexts and characteristics of the programs and projects of the region.

TABLE 9.1 Demonstration projects using research action and participatory evaluation: relevant contributions to evaluation standards

1. Education and guidance counseling that intends to be more ample, realistic, and personalized, with the development and validation of a website, and didactic modules	
Back-ground	<p>The deficiency in guidance counseling appears to be the biggest generalized abuse of adolescents in the country, leading to much frustration and unfortunate decisions by youth, and contributing to the greatly increasing nonworking youth population and violence. Making it worse has been civic education that has encouraged students to cheat on their civics graduation exams to misconstrue the need for improvements and thus for their positive contributions. It has been especially hard for today's young adults (aged 20–35) who are of a baby boom generation for which the country did not prepare. They are undereducated (only 40% have a high school diploma) and have collided with a depressed employment market</p>
Intent for target population	<ul style="list-style-type: none"> ▪ Emphasis was on the student being able to intelligently evaluate and choose his or her best prospects. ▪ Development of a website on how to choose which university or technical educational majors/specialties are most needed in the labor market and which are accessible given admissions quotas. ▪ Experimentation with guidance discourses and didactic modules to help in choosing high school and university majors and specialties, and to understand the complementary value of virtual education. ▪ Focus was on those entering their last year of high school before choosing university majors when registering for the admissions exam, those in 9th grade who often can change their category of high school, and those who must choose a category on entering 7th grade. ▪ Before giving group and individual feedback to the high school students, survey results were shared with the educators. ▪ Promotion of more forthright civic education that is honest about where youth is needed. ▪ Experimentation was undertaken largely by university graduates and professors of guidance counseling
Results	<ul style="list-style-type: none"> ▪ The survey was undertaken in 38 of Costa Rica's 81 local governments (municipalities), with 19,000 current or potential university and high school students, and the general response of the great majority was gratitude; it was undertaken in 100 mostly high schools, plus a few graduating 6th grade classes of elementary schools in low-income neighborhoods. ▪ In numerous cases, the presentation of student survey results surprised the high school educators. ▪ Some high school directors were resentful, because with this information, many of their students changed their schools to a different category; the project did, however, succeed in greatly increasing the enrollment of elementary school graduates as a result of the counseling

<p>Results</p>	<ul style="list-style-type: none"> ■ In a follow-up study of youth who had received this guidance five years earlier, most had continued with their university studies and judged the guidance as having been helpful; surprisingly, they had been very accurate in anticipating the practical problems they would encounter, such as socioeconomic background and family dynamics in pursuing their preferred prospects. ■ The elementary school graduates of very low-income neighborhoods were registered in high schools generally not of the student's preference, but rather responding to the parent's dominant criteria of security and transportation; therefore, the discourse was altered on this, telling students they could choose a different category when finishing the 9th grade. ■ Initiation of a comparative study between countries regarding civics education pinpointed the influence of the legitimization or not of militarization (given that Costa Rica has no military)
<p>Lessons learned</p>	<ul style="list-style-type: none"> ■ The target population was mostly helpful in monitoring the effects and impacts of attempts at better counseling practices. ■ This has not led to a generalized application of the better practices; attempts were made to identify the major sources of resistance to generalizing their adoption, or even to evaluation of standing procedures, particularly as this resistance was justified by referral to intellectual and organizational autonomy and/or authority that educational institutions or units insist on for themselves. ■ This resistance was very strong by universities, which are influenced by professional associations and are very slow to create new majors, specialties, and courses, and to resist the notion that guidance counseling should be more realistic—which is needed to break the tendency that the most-solicited majors are the ones leading to severe underemployment, and those majors highlighted for meeting international accreditation standards rather than for their national relevance, while there are great needs of determined specialists often in the same professions. ■ The surge of new categories of high schools caught the existing elementary and high schools unprepared, and the latter often put up obstacles against more free choice for the students. ■ The Public Education Ministry was adamant in not addressing or even wanting to hear about conflicts within its own units that are affecting student decision making, such as the conflict between civics education, apparent entrepreneurship programs, and realistic counseling; however, with a change of government, the ministry has indicated interest in taking lessons from this experience. ■ The project allowed the students a role of censorship: it was important to recognize that many students—more of middle-class than low-income neighborhoods—were emphatic about not wanting their parents involved in the evaluation and selection of a category of high school or of a university major; where possible, counseling was offered to parents, but insisting that the students could choose whether to extend the invitation to their parents

2. Identification and creation of windows of opportunity and necessity for professional and technical human resources in the zone of residence, especially of youth, to reverse their tendency to identify their best future prospects as being distant from that zone of their country, and the adaptation of formal and parallel education to prepare them for what is most relevant to where they live	
Back-ground	Local opportunities and the needs for such competencies, more as entrepreneurship than as employment, are generally ignored in community development and education plans. Most rural and many suburban communities export the resources they most need for their development, i.e., their more educated youth, even though their vocational interests may coincide with needs or potential opportunities in the local zone. Meanwhile, university and high school students undertake their required field projects, including those of community service, with very little (and at times negative) consequences for the community or target population; the discipline; or their own professional portfolios.
Intent for target population	<ul style="list-style-type: none"> ▪ The identification of the opportunities and need for professional and technical human resources in each zone, using field research and community service projects of university and high school students and community participation workshops to do this. ▪ Incidence in the curriculum to respond to those spaces and prospects of the students to respond to or develop them, and promotion of the use of virtual education. ▪ Field practice for students, e.g., regarding program evaluation, local social environmental conflicts, ecotourism opportunities, and architecture for schools or community services, all promoting the idea of conforming to a niche. ▪ Workshops and follow-up in entrepreneurship, with help in getting the required financial, technical, and institutional support. ▪ Attempts to influence governmental agencies, NGOs, community inter-institutional coordination commissions, and international cooperation to recognize the need for this, and to take lessons learned from the project. ▪ Seek ways to enable the community to have what it lacks now, which is intelligent planning, design, and evaluation of its underlying and most fundamental transformations.
Results	<ul style="list-style-type: none"> ▪ 2,000 students (fewer than expected) receiving guidance counseling in many local governments also participated in field practices related to their vocational interests; many were made more conscious of zonal needs and opportunities for their intended professional or vocational prospects, or those they are interested in. In some cases there has been immediate interest by prospective employers, such as in updated accounting and informatics for commerce of artisanal products. ▪ A few technical high schools experimented with some curricular additions, like educational ecotourism. In one local government, the lowest rated in the UN Life Quality Index, the project helped bring a new technical high school with recently created (and long needed) specialties for such schools.

<p>Results</p>	<ul style="list-style-type: none"> ■ University students of 18 majors in 6 universities participated, annexing and adapting their field projects to support the investigation and development of development needs and opportunities in the participating communities. Also, many high school students contributed to this with their community service projects. ■ In the review of Project 3 below, the subject of entrepreneurship is elaborated upon somewhat. ■ Extensive bibliographies, citing as many as 400 research reports, have been prepared in each of 4 local governments, with copies including many of the cited documents, to local authorities, and to public and school libraries.
<p>Lessons learned</p>	<ul style="list-style-type: none"> ■ The potential for much more relevant student field practice seemed quite evident, except that it is so very difficult to get the established governmental and NGO institutions to readdress their agendas and methodologies: some are set in their ways, and some are influenced by the agendas of international cooperation. Even some that highlighted youth participation in their discourse were not receptive to receiving and using the interest profiles of local youth that were offered to them. ■ There were numerous university students interested in undertaking their obligatory field activities (thesis or other graduation projects, internships, or community service) in these subjects, and the project urged them to do this in their family's residential zone. However, the universities in general are very poorly organized for such projects to be relevant in the terms the project has promoted (for the student's professional portfolio, for the community, and for the discipline). Therefore a number of initiatives were curtailed. ■ Some organizations have been influenced to take lessons from this experience, and to adapt their agendas and methodologies, but the evident need for this is being met only fragmentally. ■ The use of our bibliographies of research on local realities (a large proportion being university theses and field reports), highlighting substantial documented local empirical intelligence, has been very sporadic. Both external and local organizations prefer to embark with their preconceptions, immediate experiences, and agendas, and not be distracted by substantiated evidence of local realities.
<p style="text-align: center;">3. Social and business entrepreneurship, especially of community and agro-ecotourism projects, small businesses for low-income families, and youth initiatives</p>	
<p>Back-ground</p>	<p>Numerous studies of social entrepreneurship show that most projects have less impact than expected, such as in rural community tourism, and family business projects, including those that had received training in entrepreneurship and project development. Most are very deficient in key ways, and are not adequately prepared to deal with critical factors: most family businesses do not survive more than a few years.</p>

<p>Intent for target population</p>	<ul style="list-style-type: none"> ■ Workshops and monitoring or follow-up (including incubation), with a focus on an operative plan resolving the critical factors of the efficacy of such entrepreneurship, including segmentation of the market or beneficiaries; adjustment of the products' characteristics; realistic financial analysis; the use of information technology in marketing and transaction with the users, complementation among projects instead of overly zealous competitiveness, and in many cases, the advantages of emphasizing the social and ecological benefits of the initiative. ■ Visiting the location of each project before the workshop, in order to assure better advice from the instructors and better understanding and preparation of the entrepreneurs for what a focus on efficacy entails. ■ Periodic updating of a database with numerous (more than 60) sources for financing of these projects, with emphasis on donations. What are updated with regard to the financial sources are the themes of interest, conditions for soliciting the funds, and their approximate availability. ■ Forming of facilitators of entrepreneurship.
<p>Results</p>	<ul style="list-style-type: none"> ■ Some community entrepreneurship has been undertaken with our help, especially in cultural activity, and much family entrepreneurship as well, especially in tourism, clothing, culinary, beautician, informatics, the value-added chain of agro-products, and health service endeavors. ■ Numerous workshops have been undertaken, with a huge drop-off of those indicating interest to those with regular assistance. A big factor in the drop-off was our warning that the curriculum was demanding, requiring perseverance and patience (especially with the bureaucracy), and particularly with regard to financial analysis (although many with little formal education were able to process it with our help). ■ It appears that by far most that have regular assistance in the workshop do continue with their projects, and obtain the needed support (over 90% of the low-income women who solicited funds from a special government fund with our help, compared to less than 20% of requesters in general). ■ Very few have taken up the offer to form and give diplomas to facilitators of entrepreneurship, although demand for this is very high.
<p>Lessons learned</p>	<ul style="list-style-type: none"> ■ The project's curriculum has been criticized for being very demanding by various institutions and instructors who offer workshops and incubation in entrepreneurship. However, this exacting curriculum responds to studies indicating how few entrepreneurial projects succeed even somewhat near expectations, and the need to focus on critical factors of efficacy. Thus, an attempt has been made to emphasize more practical help than academic assistance. ■ There have been very variable differences of grades of commitment and follow-through among the communities and the families that have participated with entrepreneurial projects. The reasons for this are being analyzed by monitoring the perspectives and decisions of the entrepreneurs. There seems to be a complex set of personal, social, economic, and political factors involved.

Lessons learned	<ul style="list-style-type: none"> ■ A greatly increasing number of families who have serious health problems have participated. They are desperately seeking income to pay for private care, given the huge waiting lists in public health care services. This of course requires great attention to how to address this combination of health and financial crises. ■ Much depends on the local counterpart organizations that solicited this project, and those that either enjoin or distance themselves from entrepreneurial initiatives.
4. Attention to the increased violence in communities and schools, along with its effects; and of intrafamily violence on the psychosocial, vocational, and civic prospects for youth; and to the need to conciliate the emphasis on containing human violence with adequate preparation for disasters and emergencies	
Back-ground	<p>The major burden of responsibility in the Law of Protection of Children and Adolescents goes to educators in schools, who receive very little help with carrying it out. The programs of prevention of suicide, femicide, and abuse in general do not respond to underlying factors or statistical tendencies. The Manual of the National Emergency Commission with the Ministry of Education on how to prepare for disasters and emergencies, including brigades and simulation, is horribly unrealistic if one takes into consideration lessons learned from the earthquakes in Haiti and Mexico, as well as incidents in Costa Rica.</p>
Intent for target population	<ul style="list-style-type: none"> ■ Workshops for youth facilitators of peace and security in their schools and communities, in combination with the Ministry of Justice and Peace, Municipalities, and other agencies dedicated to the subject. ■ Advice to school student commissions dedicated to the theme. ■ Workshops for educators on how to assist adolescents traumatized by violence, indicating that there is a government agency to help them in case of reprisals. ■ To create a commission on how to mitigate the huge arrival of so many arms, drugs, fugitives (including pedophiles) into the country, and the impact of the romanticizing of violence, most of this from the United States. ■ On-site evaluation of how to conciliate the great attention to containing human violence, by enclosing the students and educators in barbwire fencing, small gates, and guards (who are often off on errands), and lack of the open spaces not under roofs and ceilings, with realistic preparation for the mass movement of students, educators, rescuers, hysterical parents, and water, sediments, and lack of electricity in case of a disaster such as an earthquake, fire or huge storm.
Results	<ul style="list-style-type: none"> ■ Only two local governments responded to the concern about the contradiction between measures of containing human violence and preparation for disasters and emergencies, facilitating an investigation, which made recommendations to 17 schools. However, technical solutions for key problems, such as economical remote locks that the guards can handle at a distance and that are not dependent on the electrical system have not been found.

<p>Results</p>	<ul style="list-style-type: none"> ■ In two other local governments, there was a surprising regularity of attendance at our workshops for 125 youth facilitators of peace and security in their schools and communities (including students from 25 high schools), given the schedule of 6 all-day Saturday sessions, and the distances to travel that were required. Through gaining confidence with the students, and interviewing them, it was understood that the great majority were at least partially motivated because they have violent internal family situations, and most had not shared that information with anyone. ■ Naturally, they contributed considerably to penetrating the world of violence in which they are living, and are observing. ■ The students gained confidence by their preparing 25 videos of their own artistic creations, including themes not presented in the conferences of experts and authorities. Their artistic renditions have been circulated to many schools and local child protection committees. ■ Advice was given to municipal commissions, educators, and student commissions on their respective work plans regarding violence, and help in articulating with local agencies. ■ Several high schools enabled the training of educators on how to attend to adolescents traumatized by violence. ■ The government agency in charge of helping endangered victims and witnesses did not show an interest in an offer to create a video about their services to circulate among educators. ■ This project was curtailed abruptly, as the Ministry of Education created new protocols, insisting that they were to be “supreme and exclusive,” although they were not adapted to the realities discovered in the demonstration projects.
<p>Lessons learned</p>	<ul style="list-style-type: none"> ■ Organized commissions in two municipalities facilitated the project’s activity of the youth in favor of more peace and security in the community, and officials in two other municipalities facilitated inspections on how to prepare schools better, with periodic evaluation of what was undertaken, until the project was curtailed abruptly. The Public Education Ministry is now, several years later, reviewing what happened in this regard. ■ The last few years we have attended 11 forums on dealing with violence in communities and schools in Costa Rica, with more than 65 conferences by international experts and national counterparts of executing agencies, of which very few shared lessons of the realities of their field experiences. By far most made propaganda for what was and will be intended. Several key themes or factors in the game were never mentioned in the conferences, and only in a few cases was there opportunity for discussion. ■ The governmental and university organizations dedicated to preparation for disasters showed no interest in the problem of conflict of measures regarding violence that have put so many students and educators in jeopardy. This is now being undertaken, but the criteria for selecting model schools are very deficient and impractical; and the models have serious defects. Hopefully, this new program will take lessons from what was undertaken by the project.

5. Taking lessons from exceptional patients, with due attention to critical factors that are receiving inadequate attention	
Back-ground	<p>Many recent panel studies of clinical trials of oncological patients in different countries show that 5–10% of patients survive more than three times the average life span of all patients receiving treatment. There is a great need to draw lessons from the experience of these and other exceptional patients.</p>
Intent for target population	<ul style="list-style-type: none"> ▪ Anthologies of biographies of exceptional patients to see what antecedents, conditions, coping, and complementary treatments they have or have adopted. ▪ Special attention to the huge challenges for the family caregiver, internal familial tensions, impact on family children and adolescents, medical and neuropsychological services, support groups, and implicit suicide chosen by so many ill people who are suffering more financially, or for other reasons, than physically or emotionally. ▪ Wide circulation of these lessons, to enhance medical and neuropsychological attention beyond just following epidemiological tendencies, to help transcend those tendencies; this can take the form of textbooks, manuals for patients and family caregivers, protocols for medical personnel and neuropsychologists, indicators for researchers, and brochures for children and adolescents. ▪ Helping with comparative (and hopefully also international) studies of R&D on these patients, with an emphasis on M&E of the impact of agency programs for patients and family caregivers.
Results	<ul style="list-style-type: none"> ▪ Intensive research is in process in very extraordinary cases to serve as prototypes for more extensive studies. ▪ Preliminary findings have been shared with medical personnel and neuropsychologists, family caregivers of dementia and palliative care patients, and with support groups—e.g., for the 115 Costa Rican cancer patients who suffered overirradiation in their cobalt treatments. ▪ Some exceptional university programs and NGOs in public health do show interest in what is being discovered, and in seeing how to adjust their agendas and procedures accordingly, especially with family caregivers. ▪ Key challenges continue to be treated very marginally or exceptionally for the seriously ill, including exceptional patients who are also being evaluated: e.g., children in families with a medical crisis or premature death of a very ill sibling or parent; in general, intrafamily relationships that are so often tense and conflictive, affecting both the patient and the key family caregiver. ▪ Social health agencies and programs interested in impact M&E of their programs are being helped to engage university students undertaking their fieldwork or internships, with the intention that when feasible they can elaborate on and test badly needed protocols.

Lessons learned	<ul style="list-style-type: none"> ■ The experience of 115 cancer patients who in 1996 suffered overradiation is a prelude to what is evident in Costa Rica and even more elsewhere, such as the United States: that legal, paradigmatic, and ideological currents can greatly impede due attention to factors that can explain the prolonging and quality of life for exceptional patients. ■ The emphasis has been on indemnification for malpractice palliative care, ideological debates regarding euthanasia, and the usual benchmarks in clinical trials, at the expense of due attention to the possibility of prolonging the length and quality of patients' lives. ■ A key problem for evaluation can be the chaos or "snarl" in handling patient medical records, including legal handling. ■ Family caregivers are more receptive to taking lessons from exceptional patients, but they are imbued with very haphazard support for what they have to resolve, and ironically the meager attention they do receive focuses almost exclusively on some of the onerous tasks before them, and not on the personal benefits they often derive from caregiving; this is evident in the most frequently used questionnaires and protocols for caregivers. ■ The most positive impact has been with patients, but the impact is limited by the enclosed perspectives of the general discourse on this subject, and that of professional and informal caregivers. ■ There is attention to an incipient movement, especially in Europe, to alter the perspective of medical personnel and neuropsychologists on this matter, e.g., that palliative care can be given early on, and may assure more prolongation as well as quality of life, not just preparation for a more serene death.
6. Knowledge management for older adults	
Back-ground	<p>Costa Rica's legislation is rather unique in recognizing the importance of two particular interests with regard to older adults: to enhance, adapt, and take advantage of their accumulated empirical intelligence; and to have a direct role in the design and evaluation of programs and projects for their generation. However, we see very little application of these principles.</p>
Intent for target population	<ul style="list-style-type: none"> ■ Guidance on recapitulation of his or her accumulated intelligence, and how to transmit the lessons and benefits of that knowledge to succeeding generations. ■ Voice and vote in the design and evaluation of the programs and projects intended to enhance the older adult's quality of life.
Results	<ul style="list-style-type: none"> ■ There has been spotty acceptance among some public and private agencies in helping foster and organize this initiative, but overall there have been many obstacles. ■ There has been much less progress than hoped for with these intentions, and thus with the evaluation of the experience. ■ In one suburban zone, an older adult association is strongly committed to pursuing these intentions; which offers some hope.

Lessons learned	<ul style="list-style-type: none"> ▪ Older adults, including the more educated, believe that they are to be retired, respected, entertained, and perhaps educated in some subjects, but not with the type of activity intended in this project. ▪ To our surprise, much more than with professionals attending to adolescents and youth, where adult-centrism was expected (in projects 1–4 above), those attending to older adults generally were adamantly against these objectives, possibly fearing a conflict between the need for caregiving and program management. ▪ The older adult is beset with stigmas held by others, and even with self-adoption of such stigma.
7. Effects on communities and users of transportation and communications public service investments in buffer, forest, plantation, and mining zones	
Back-ground	<p>There is a strong sense that these investments enhance transitory exploitation and not sustainable development of these zones, contributing linking to socio-environmental conflicts; great distortions in the added value chain of local products; and higher user costs affecting all sectors. This is an example of a problem or conflict resolution apparently being a predictable precedent of others.</p>
Intent for target population	<ul style="list-style-type: none"> ▪ To know the role of transportation and communications in the evolution of community and family economic sustenance and survival strategies. ▪ To study the options and the predominant criteria of investments in these sectors. ▪ To know the role of the predominant models for these investments and protocols by academics and international cooperation.
Results	<ul style="list-style-type: none"> ▪ Initial studies of Central Appalachia in the United States, and central and southern Costa Rica. ▪ Probing of other case studies and a historical overview of the application of international cooperation models in Latin America. ▪ Preliminary results indicate historical adoption of similar investment models of earlier private investors and later public agencies. ▪ The highly influential evolution (really, several huge evolutions) of the globalization strategy of transnational corporations has not been well addressed in economic development studies of Latin America.
Lessons learned	<ul style="list-style-type: none"> ▪ The significance of this for the community and families who live in it has been largely overlooked by economic historians and anthropologists. ▪ Interviews with older residents of these zones give inconsistent versions of community and family history; there is a need for more creative methods, such as theatrical representations, of that reality. ▪ This appears to vouch for the great need in Central America of R&D-focused university majors in transportation economics.

The standards for Latin America are also a guide, a way forward in the design and execution of evaluations that will provide greater knowledge, clarity, and understanding of the contexts in which policies, programs, and projects are being developed, and will thus improve evaluation processes.

REFERENCES

- ICAP and UNDP (Central American Public Administration Institute and United Nations Development Programme). 2003. "Perfil de Cooperante y del Recipiente Responsables." Vol. 3. Guía de Gerencia de la Cooperación Internacional.
- Crowther, Warren. 1999. *Manual de Investigación-Acción para la Evaluación Científica en el Ámbito Administrativo*. San José, Costa Rica. Editorial Universidad Estatal a Distancia.
- Rodríguez Bilella, Pablo D., Sergio Martinic Valencia, Soberón Alvarez, Sarah D. Kilier, Ana Luisa Guzmán Hernández, and Esteban Tapella. 2016. *Evaluation Standards for Latin America and the Caribbean*. 1st ed. Buenos Aires: Akian Graphic Publishing.
- Stockmann, Reinhard. 2011. *Manual de Evaluación: una guía práctica de procedimientos*. Universidad de Costa Rica, Editorial de la Universidad de Costa Rica, Costa Rica.
- UNICEF Evaluation Office. 2002. "Children Participating in Research, Monitoring and Evaluation (M&E) Ethics and Your Responsibilities as a Manager." *Evaluation Technical Notes* No. 2. (Spanish translation by Priscilla Hurtado and María Isabel Araya Tristán.)
- UNICEF Costa Rica. 2000. "Código de Ética para las ONGs que participan en proyectos." Internal document.

Chapter 10

The Role of Monitoring and Evaluation in the MENA Region, with a Focus on the Arab Uprising Countries

Sonia Ben Jaafar and Awny Amer

Abstract. *The demands of monitoring and evaluation (M&E) for the Sustainable Development Goals in the Middle East and North Africa region are particularly challenging. The diversity of the 22 countries; a large and youthful population; unprecedented political transitions; and a variety of conflicts and humanitarian crises have created a cluster of complex development needs. Development aid activity in the region is now operating in a landscape of high accountability demands and traditionally poor governance practices. This chapter explores the way these regional transitions are cultivating a new paradigm that promotes national capacities and country M&E systems in which local M&E professionals are valued and supported in the development of their own countries and region.*

The Arab region is comprised of 22 countries in Northern Africa, the Levant, and the Arabian Peninsula. This region, which is known as the Middle East and North Africa (MENA) region, has a population of approximately 350 million, the majority of which is under the age of 25.¹ The current challenges affecting this region are partially tied to an unprecedented wave of political transition that has swept the region since 2010, coupled with an increasing number of conflicts. According to the United Nations Development Programme (UNDP):²

Weak social, political and administrative accountability mechanisms and politically oriented socioeconomic planning models have resulted in the neglect of large parts of the population. These nations face the challenge of forming new, accountable governments that reflect popular aspirations.

THE ARAB UPRISING

The Arab Uprising refers to a series of antigovernment protests, uprisings, and armed rebellions across the Middle East that surfaced in 2010. By 2012, the rulers in Tunisia, Egypt, Libya, and Yemen had been pushed out of power; civil uprisings had erupted in Bahrain and Syria; there were major protests in Algeria, Iraq, Jordan, Kuwait, Morocco, and Sudan; and minor protests had occurred in Djibouti, Mauritania, Oman, Palestine, Saudi Arabia, and Western Sahara (Smith 2016). This revolutionary atmosphere in the Arab region was tempered in the Gulf states, where a more reformist approach took root (Abdalla 2012). The Gulf monarchies demonstrated the ability to adapt to regional shifts and to address internal issues with policy measures, by using their “oil wealth, historical legitimacy, Bedouin culture, demographic scarcity, extensive security services, patriarchal regimes, and the absence of an opposition and political parties” (Abdalla 2012, 30).

Moreover, according to UNDP’s 2015 *Human Development Report*, Bahrain, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates are ranked among the top countries in the Arab region in terms of average income per capita (UNDP 2015) and the economic competitiveness index,³ which creates a certain level of comfort for their people. For example, the United Arab Emirates leads the Arab states, and is considered globally as one of the happiest countries: in the 2015 *World Happiness Report*, it is given special mention as an example of a country in which well-being has been made a central tenet of the design and delivery of the national agenda (Heliwell, Layard, and Sacha 2015).

¹ Source: United Nations Development Programme, <http://www.arabstates.undp.org/content/rbas/en/home/regioninfo.html#Introduction>.

² Ibid.

³ World Economic Forum, “Competitiveness Rankings,” <http://reports.weforum.org/global-competitiveness-report-2015-2016/competitiveness-rankings/>.

However, for the purposes of this chapter, we will discuss the challenges with an emphasis on the majority of the region, where there has been revolutionary upheaval and a large impact on development. We focus on the impact of more than 11 million people forced from their homes in Syria;⁴ an estimated 7 million internally displaced people within Syria; and more than 4 million who have fled Syria as refugees, a large portion of them landing in Jordan, Lebanon, and Turkey. We focus on the impact of an initially peaceful uprising in Libya that quickly became an armed conflict with Western military intervention. According to Amnesty International,⁵ once Libya had become deeply divided, the internal conflict has caused civilians to live in a constant state of threat, with nearly 2.5 million people in need of humanitarian assistance, clean water, sanitation, and food since 2014. In this chapter, we focus on the areas in crisis, which has drawn international attention and aid, because the potential to meet the Sustainable Development Goal (SDG) targets in the region is increasingly bleak.

DEMANDING ACCOUNTABILITY

The key driver in this unprecedented regional call for change is the call for greater accountability from the government to the people. This root demand has thrust the role of monitoring and evaluation (M&E) to a new level of significance, giving it greater value for all stakeholders. As M&E and accountability have been becoming more important to the local populations, donors are more interested in the *effectiveness* of aid in the region. The SDGs that are key priorities in MENA countries focus on ongoing and intersecting issues of poverty, hunger, health, education, energy, equity, economic downturns, climate, peace, and stability in the region.

As aid funding increased in a progressively complex and challenging region during the aftermath of the Iraq War, so did attention to developing specific mechanisms to ensure aid effectiveness. This promoted solutions like the application of results-based management to program, thematic, and sectoral evaluations, rather than simple project evaluations (UNDG 2011). Solutions such as country-level evaluations that consider coordinating the efforts of multiple donors with joint criteria will contribute to ensuring the effectiveness of the evaluation processes (Baradei, Abdelhamid, and Wally 2014). Although this shift is not restricted to the Middle East and North Africa, the region has grave challenges in adopting these solutions because of traditional social, political, and economic practices that are unaccustomed to the three key pillars of results-based management: accountability, national ownership, and inclusiveness (UNEG 2011), pillars that happen to align with the reverberations of the Arab Uprising.

⁴ Source: Syrian Refugees website, "A Snapshot of the Crisis in the Middle East and Europe," <http://syrianrefugees.eu/>.

⁵ Source: Amnesty International website, "The Arab Spring' Five Years On," <https://www.amnesty.org/en/latest/campaigns/2016/01/arab-spring-five-years-on/>.

Given this reality, the M&E community in the region has been increasingly focused on accountability, national ownership, and inclusiveness. Some of the issues of core importance to the M&E community follow:

- **Accountability** as it pertains to the extent of the relationship between the impact of aid effectiveness and the level of effort from implementing and donor organizations; and the value added of M&E processes in terms of corrective program actions and policy changes
- **Inclusiveness** as it pertains to the role of the beneficiaries of aid efforts in relation to implementing partners; the traditional North-South paradigm, in which development interventions and M&E are conducted primarily for donors; and the lack of shared accrued knowledge derived from M&E that adds to the future independence of local actors and organizations
- **National ownership** as it pertains to the extent of local M&E capacity and professionalism; and the extent of involvement of local leadership in assessing the effectiveness of the aid received

These areas of focus align tightly with evaluation of the SDG plans and programs that necessitate the involvement of national efforts to cultivate appropriate evaluation capabilities at all levels and across all stakeholders (El-Saddik et al. 2016). It is noteworthy that although accountability is a shared value across the globe, with good governance as a core element, the Arab region lagged behind other regions in most governance indicators in 2015 (UN 2013). Given that good governance is not only essential for accountability, but is also a gateway to inclusion and national ownership, M&E efforts in the region are arduous and highly political. Local M&E professionals work within very challenging parameters in their attempts to change the landscape on the path to meeting the SDG targets.

An illustrative example of the limited role of M&E in the region is the need for the Taqueem Initiative, which was established in 2009 by a partnership between the International Labour Organization, Silatech, the World Bank, the Jacobs Foundation, and the Swedish International Development Foundation. Taqueem was specifically created to support youth employment policy makers and practitioners in enhancing the M&E of their programs (ILO 2009).

The lack of evaluation information on youth employment programs in the region is telling, given that it is mostly populated by youth,⁶ and that the youth employment challenge is a critical issue that threatens the already fragile economic and political state of many countries in the region. Creating opportunities for Arab youth has long been a leading policy priority for international organizations, governments, nongovernmental organizations, and social entrepreneurs. Traditional and nontraditional programs, projects,

⁶ Source: UNDP, <http://www.arabstates.undp.org/content/rbas/en/home/regioninfo.html#Introduction>.

initiatives, and partnerships have been thrust upon the region by both local and foreign organizations. The Taqeem Initiative targets these particular solutions and helps local organizations drive quality M&E while creating a regional database of evaluation information that will help policy decisions concerned with solving the youth employment challenge.

Although institutionalization of the Taqeem Initiative is laudable, it is regrettable that the culture of the region did not already have the demand and supply of M&E for these solutions to meet a vital regional challenge. The reality of the situation is well summarized in a recent study of the state of M&E in post-revolutionary Egypt, in which the authors highlight the fact that the government urgently needs effective support “to demonstrate and measure the results of each policy” (Baradei, Abdelhamid, and Wally 2014).

THE CHALLENGES OF EVALUATING RESULTS FOR DEVELOPMENT

The regional M&E community has been sharing knowledge to identify key issues that contribute to promoting the demand for the evaluation of development results in response to a rising public demand. A series of strategic discussions and debates across MENA countries have taken place to better appreciate the current situation of M&E processes and practices in the region (El Kabbag 2011). The following key challenges have been identified as barriers to normalizing supportive M&E in the region.

A deterring culture of evaluation (El Kabbag 2011). In most MENA countries evaluation is tightly coupled with undesirable beliefs about the impact of the results. In this context, evaluation has either no influence at all, or grave consequences for local stakeholders and projects. Evaluation is believed to be restricted to audits and financial reviews that demonstrate inefficiencies or corruption, or public reviews that expose fraud. It is also believed that it is only relevant to donors for the purpose of satisfying bureaucratic requirements; that it is conducted apart from the beneficiary communities; and that it is tied solely to outputs, with no regard for outcomes and impact, and all of the other traditional high-stakes and potentially harmful ways evaluation has been used in the past.

The status quo is the norm. Most of the organizations in MENA have a culture of following evaluation processes that are stagnant within the organization. This static reality is in contrast to an evaluation approach that focuses on a dynamic and holistic perspective designed to transform the organization for the better.

National capacities are not ready for change. Given that the MENA region has not had a strong historical integration of M&E into their public governance, it is not surprising that there is a less than adequate national capacity in many of the countries, at both the governmental and civil society levels. In particular, evaluators in the region generally do not have ready access to accredited M&E training, and have gaps in some of their knowledge and

skills. For example, there are gaps in adopting minimum standards of ethics during the evaluation process. There is also a gap in knowledge on how to use system-level data, because many governments have incomplete, inaccurate, insufficient or dated data.

Lack of good governance affecting good evaluation. As previously stated, the Arab region has one of the poorest governance records globally. When we consider the indicators of good governance, such as voice, accountability, and governmental effectiveness, it is evident that poor governance will have a significant influence on evaluation practices (Kaufmann, Kraay, and Mastruzzi 2010). This situation is part of the overall socio-political context that is being called into question with the Arab Uprising.

A non-enabling environment for developing evidence-based policy. Evidence-based policy is rooted in responsible government, which has become a mantra in the post-Arab Uprising MENA region. Evidence-based policy is expected to reduce wasteful spending, expand innovative programs, and strengthen accountability (Kaufmann, Kraay, and Mastruzzi 2010). Part of the framework for this approach, which enables governments to make better choices, is program assessment, outcome monitoring, and targeted evaluation. Regrettably, a lack of good governance is often tied to a lack of evidence-based policy. The situation in the Arab region is such that the tracking of progress at the policy level for SDGs and Millennium Development Goals has been, and continues to be, a key challenge. This means that in addition to a low supply, there is a low demand for M&E.

PROFESSIONALIZATION OF M&E IN THE MENA REGION

The challenges presented in the previous section are directly tied to the systematically poor governance that created the conditions for the Arab Uprising. M&E supply and demand have been less significant in the past, but the uprising, and the subsequent attention to SDGs have exposed an urgent need for M&E. Local governments and the international community are now promoting accountability and unveiling a substantive demand. At the same time, the M&E communities of professionals have been working to create an enabling environment for M&E. The regional professionalization of M&E specifically promotes national ownership and inclusion, and activates an evolved accountability that goes beyond the conducting of external audits to appease donor distrust (Segone 2009).

The increasingly high demand for M&E professionals in the MENA region is an issue, especially since quality M&E requires an investment of time and resources in order to plan, collect data, and report appropriately. Increasing efforts to professionalize the field are leading to improved evaluations and better evidence. This increases the demand for more and better evidence, which in turn stimulates further improvements in professionalism. However, program and operational funds in organizations in this region often neglect to budget for appropriate M&E, for various reasons. The key issue, a lack of data culture, which makes for a situation in which it is difficult to

cultivate M&E professionals, has been addressed with multiple efforts by international and local organizations.

In particular, the launch of the Middle East and North Africa Evaluators Network (EvalMENA) in 2012 galvanized M&E professionals at the regional level. EvalMENA is an informal network of stakeholders who are dedicated to professionalizing M&E in the region. It aims to promote and strengthen M&E culture and practices in all the MENA countries by encouraging country-level professional M&E associations, and by offering a networking platform for thoughtful debate about local issues and solutions.

The work of EvalMENA is practical and addresses local issues. The success of the organization is largely due to the fact that inclusiveness was—and continues to be—a fundamental building block of the organization and all the practices of EvalMENA. This success includes lobbying for M&E, networking events, capacity building, creating new evaluation associations, and professionalization efforts. The inception of the organization was a research and development project championed by the Environment and Sustainable Development unit of the American University of Beirut, with technical and financial support provided by the International Development Research Centre in Canada (IDRC). However, even with donor funding and appropriate supportive guidance from IDRC, the leadership and working teams have always been local Arab professionals. Since its inception in 2008, EvalMENA has managed to achieve the following (Moussa 2015).

Bringing MENA evaluators together, in the region and around the world.

In 2013, the membership was made up of 146 evaluators from 23 countries. By 2015, the membership had grown to 360 members from 38 countries. Most of the members were from Egypt, Jordan, Lebanon, Morocco, Palestine, and Tunisia, with a small minority from North America. Currently, the membership is over 500 members, indicating a growing interest and support for the vision of EvalMENA.

Promoting and supporting seven new national evaluation associations.

National ownership and local inclusion has been a top priority of EvalMENA since the inception of the network. The formalization and maturation of the Moroccan Evaluation Association (MEA) has been an important catalyst in the region. MEA has now matured to the point where its work with policy makers in Morocco is allowing them to take an important seat at the same table with the highest level of government. The success of MEA has proved to be a catalyst in the formalization of six more new networks: the Egyptian Research and Education Network (EREN) in 2012; the Palestinian Evaluation Association (PEA) in 2013; the Jordan Evaluation Association (EvalJordan) in 2014; the Tunisian Evaluation Association (RTE) in 2014; the Lebanese Evaluation Association (LebEval) in 2014; and the Egyptian Development Evaluation Network (EgDEval) in 2015.

All of these new associations are now actively contributing to the leadership and management of EvalMENA, and are promoting its vision at the national level.

Organizing five annual regional conferences between 2012-16. The Arab culture places a high value on face-to-face networking. Given the critical importance of bringing M&E professionals together to share and create knowledge that is localized for the region, it was—and is—important to hold local conferences. As the membership and national ownership has grown, EvalMENA has shared and distributed leadership and responsibilities among the participating countries, while offering them ongoing support. As table 10.1 indicates, the national organizations have taken ownership of the regional vision. This is especially impressive after the initial IDRC funding ended, as the national and regional organizations collaborated to find other funding sources to continue the work.

Promoting MENA to an international audience through active leadership in the global evaluation movement. The context within the MENA region is becoming increasingly challenging, with humanitarian crises becoming the norm. The stunted progress in many Arab countries is cultivating high-risk zones. This creates a two-pronged problem with respect to local capacities. First, local M&E professionals are not plentiful in the region, and donors are more comfortable with M&E professionals who have more experience and formal training. Second, local M&E professionals are in greater demand, because foreign M&E officers are less and less able to access the zones where information is most needed.

EvalMENA has been working strategically to achieve recognition for the region's M&E professionals from the international community. Given that most donors are part of the international community, this recognition is important for donor trust of local capacity.

As regional annual conferences have gained traction in the region, M&E professionals have gained increasing access, awareness, and encouragement

TABLE 10.1 EvalMENA conferences

Year	Location	Key donors
2012	Lebanon	International Development Research Center (IDRC)
2013	Lebanon	International Development Research Center (IDRC)
2014	Jordan	International Development Research Center (IDRC)
2015	Egypt	UNICEF Regional Office for MENA (MENARO) IDRC IOCE/EvalPartners International Initiative for Impact Evaluation (3ie) Islamic Corporation for the Development of the Private Sector (ICD) Environment and Sustainable Development Unit of the American University of Beirut (AUB-ESDU)
2016	Tunis	Tunisian Evaluation Network (TEN) International Initiative for Impact Evaluation (3ie) UN Women-Regional office IOCE/EvalPartners

to be actively involved in the global evaluation movement. For example, active EvalMENA members have become executive board members of the International Development Evaluation Association (IDEAS), the International Organization for Cooperation in Evaluation (IOCE), the Africa Gender and Development Evaluators Network (AGDEN), and the African Evaluation Association (AfrEA), including the AfrEA presidency in 2012 and the IOCE presidency and EvalPartners cochair for 2015–16. These leadership roles have demonstrated to many members of the international community that high-level M&E professionals are present, and ready to contribute to the work in the MENA region. It has also shown Arab M&E officers that there are opportunities to grow professionally, and to engage with global networks.

Launching of the first online training course on development evaluation in Arabic. The first online training course on development evaluation in Arabic was launched in 2014 on the “My M&E” portal. This is a significant accomplishment because it is an Arabic course that is listed on a high-profile site along with courses from UNICEF and UN Women. This course boasts a credibility that is valued by donors and international agencies. Since it was launched, 276 Arab-speaking individuals with an interest in learning about M&E have successfully completed the course. It is noteworthy that many of the participants are from Iraq, Sudan, Syria, and Yemen, countries where there is no national network for evaluators. This online course helps Arabs who have limited local access to opportunities learn and share knowledge about development evaluation overcome an important barrier to joining the M&E community. This is particularly important as online learning, as an information and communications technology (ICT)–based solution, becomes increasingly vital for promoting access to learning in conflict and poverty-affected areas. The course continues to be accessed by Arabs who are hopeful about further developing their M&E knowledge, skills, and professional network.⁷

Facilitating South-South collaboration on evaluation within and beyond the MENA region. As we move into a new paradigm that values and promotes inclusion and national ownership in M&E, it is essential that there is a shift from the traditional North-South donor-recipient mentality to a shared values and joint work mentality. South-South collaborations promote the notion that all parties to the partnership are equally valuable, and that they aim to achieve a shared goal. In strategically bringing together and supporting so many actors in the region, EvalMENA has cultivated a culture for knowledge sharing and creation. In addition to the EvalMENA-sponsored national workshops and events organized in Egypt, Jordan, Lebanon, and Morocco, three projects won the 2014 Peer-to-Peer small grants from IOCE:

⁷The course is available at <http://www.mymande.org/elearning/course-details/6>.

- **Media Promoting Evaluation Culture in MENA.** A joint partnership between EREN, the Faculty of Communication at Cairo University, and Environment and Sustainable Development Unit /EvalMENA;
- **Integrating Evaluation in Legislative Bodies** A joint partnership between EREN, the Secretariat of the Egyptian Parliament, and the Parliamentary Forum for Development Evaluation in South Asia; and
- **The Evaluation Database Enhancement Project.** A joint partnership between the Palestinian Evaluation Association and the Jordan Development Evaluation Association.

All of these accomplishments in professionalizing M&E in the region would be notable at any time. But it is particularly impressive and relevant given that all of these activities took place in the aftermath of the Arab Uprising, one of the most turbulent periods in the recent history of the region. The instability of the region necessitates rethinking M&E approaches in general, but especially in areas where there are humanitarian emergencies. This is particularly true where there is an unprecedented humanitarian crisis with predictions for worse to come, as is the case in this region. According to the United Nations High Commission for Refugees, the conflict and violence in Iraq and Yemen have displaced 4.5 and 2.18 million people, respectively; and the conflict in Syria has displaced 4.8 million people who are seeking safety in Jordan, Lebanon, Iraq, and beyond.⁸

The commitment of the national M&E communities to develop national evaluation capacities for the SDGs was exemplified by Egypt's willingness to conduct a voluntary review of the SDGs at the United Nations High Level Political Forum on Sustainable Development in 2016. More than 60 representatives from the private sector, civil society organizations, foundations, academia, youth, special interest groups (women, environment, startups, etc.), and development partners were invited by the government of Egypt to discuss the road map for implementing and monitoring the SDGs. The meeting was hosted by the Ministry of International Cooperation, and co-organized by UNDP and the World Bank, as part of their joint effort to raise awareness about the SDGs in Egypt. This group focused on Egypt Vision 2030, the national sustainable development strategy.

In a spirit of accountability, Egypt volunteered to report on their efforts to achieve the agenda: they reported that engaging with local interest groups is pivotal to building ownership of the SDGs; capitalizing on local efforts for knowledge and resources; and increasing mutual accountability of national development results.⁹ The practical outcome of the report is a series of elements that need to be addressed through collective action. These are as follows:

⁸ Source: United Nations High Commission for Refugees website, <http://www.unhcr.org/en-us/syria-emergency.html>.

⁹ Source: UNDP website, "Building Ownership of the Sustainable Development Goals in Egypt," <http://www.eg.undp.org/content/egypt/en/home/presscenter/articles/2016/june/building-ownership-of-the-sustainable-development-goals-in-egypt.html>.

- **Tackle the data gap, and map national status.** The local groups expressed their readiness to support the monitoring and follow-up of the SDGs. They identified the need to set baselines for all SDG goals and targets.
- **Integrate the implementation efforts of local groups led by national authorities.** A big challenge is to integrate the efforts of government bodies and ministries for the implementation of the SDGs with other nonstate actors. Terms of reference will be developed for different goals where a multidisciplinary working group composed of government and key group representatives will support the implementation of the SDGs.
- **Build partnerships with nonstate actors, especially the civil society sector.** In parallel with the review work, Egypt Vision 2030 was formulated through an extensive nationwide consultation process. The 2030 agenda provides a platform for government and nonstate actors to initiate discussions around common areas of work, and to complement their efforts to achieve national priorities.

As in Egypt, all of the countries in the region are promoting greater local capacity and ownership of M&E as a key element in strengthening M&E practices in connection with the SDGs. The notion that local M&E talent should be used is obvious within a modern paradigm where inclusion and national ownership are core values. For those organizations that have yet to shift into this approach, there is a more practical reason to leverage local M&E capacities. According to a study released in 2015 on M&E practices during humanitarian emergencies (Jansury et al. 2015), M&E is a means for international organizations to use in addressing issues of access and security in complex emergencies. Organizations can choose to either work through local partners (nongovernmental organizations), or to employ local staff.

Although there are challenges to including local M&E professionals, the benefits have been shown to be important. For example, in 2010 the United States Agency for International Development (USAID) mission implemented the Yemen Monitoring and Evaluation Project. This project used third-party local partners to provide on-the-ground performance monitoring, verification, and evaluation of USAID activities. They were able to successfully identify problems in the quality of some of the Community Livelihood Project's rehabilitation activities, and of goods delivered. This was especially important as security deteriorated after the Arab Uprising, and U.S. employees could no longer access some regions of the country to monitor and identify problems with project activities (Office of Inspector General 2015).

These are the kinds of success stories that demonstrate that local M&E professionals can, and should, be involved in the work of foreign aid projects. It is unfortunate that this is not a normative practice despite the discourse concerning local engagement among many donors and foreign aid organizations. For example, the USAID Office of Inspector General conducted a survey to identify the challenges USAID faced during the early transition period (Q4 2010–Q3 2014) in Egypt, Libya, Tunisia, and Yemen. One of their conclusions

was a recommendation to employ third-party monitors in transitional situations because they “can help an office gain access to regions of a country that become inaccessible to U.S. direct hires or when U.S. personnel are ordered to evacuate” (Office of Inspector General 2015).

Moving beyond the “absolute necessity” argument of international organizations delivering foreign aid, local M&E professional communities are shifting the role of M&E within their local governance systems. For example, PEA has offered training on the evaluation of humanitarian programs that is focused on evaluation design and methodology; facilitated national round-table discussions on the evaluation of SDGs; and collaborated with members of the Palestinian Legislative Council to systematize M&E activities within government initiatives. Another example is EREN, which has co-offered seminars on evaluation for improving governance practices with Plan International, and workshops with government officials on results-based management with UNICEF. They have also been working with local governments’ policy briefs, and through a national conference on country-led M&E have focused on promoting transparency and efficiency with policy makers.

These efforts have been matched with those of other local organizations committed to including M&E professionals on staff despite the supply challenge. For example, a recent study in Egypt found that most organizations involved with development M&E provide training internally, in which “much of the training conducted is done through on-the-job learning” (Baradei, Abdelhamid, and Wally 2014). In addition, local organizations face a high turnover in the M&E role, in part because qualified M&E officers are in short supply and thus often highly sought after by other organizations (ILO 2009). This often means that when local organizations do dedicate the time and budget to training program officers to become M&E officers, other organizations—often international ones with greater resources than local ones—poach talented and trained program officers once they have the experience and background to conduct M&E work in the region (Boitnott 2015).

CONCLUSION

Well-intentioned external parties often treat the MENA region as a monolithic entity. But the countries within the MENA region are very diverse. UNDP identifies four distinct groups of countries within MENA: the Mashreq and Maghreb countries, the countries of the Gulf Cooperation Council, and the least developed countries (UN 2013). International organizations that do not differentiate the needs within the region learn their lesson after implementation issues have been identified. For example, in the evaluation of the Media Cooperation under the Danish Arab Partnership Programme (2005–12), which was commissioned by the Danish Ministry of Foreign Affairs, it was found that the cultural context differs in Jordan, Lebanon, Morocco, Syria, and Tunisia, creating conditions in which a uniform regional approach across all Media Cooperation Programme themes is not conducive to achieving the best results (Ministry of Foreign Affairs of Denmark 2013). In this regard, it is worth saying that commissioners of evaluations tend to lay the responsibility for the findings with the evaluators, not with themselves.

The inclusion and national ownership movement is essential for finding sustainable solutions for one of the most volatile youthful parts of the world. The paradigm shift in terms of development in the MENA region is tied to the current state of development in general. Cultivating accountability, inclusiveness, and national ownership is a challenge in a region where there is a continuing war, and a growing humanitarian crisis. (More than 5 million people have fled Syria since 2011, seeking safety in neighboring countries, and millions more are displaced inside Syria.¹⁰) The unprecedented challenges in the region are drawing international attention and funding, with various solutions to help minimize the loss of life, hope, and health, and to reduce the conditions of indignity for many people. These solutions are being constructed as aid provided with a traditional approach, which has yet to move beyond external accountability, and to value inclusiveness and national ownership. At the same time, measuring aid effectiveness in hazardous areas is opening new opportunities for local M&E professionals to gain experience, trust, and training from international agencies simply because these are the M&E people who are on the ground with access to the conflict areas (Jansury et al. 2015).

The evaluation efforts tied to the SDGs in MENA are increasingly aligned with the key principles of accountability, inclusiveness, and national ownership. This triad cultivates accountability for sustainable development when public policies are subject to local evaluations whose purpose is to ensure the best solutions for local issues. This approach means explicitly addressing the data gap and mapping national status along the SDG indicators; integrating the efforts of major groups involved in implementation, led by national authorities; and building partnerships with nonstate actors, especially with civil society. This can only be achieved when local M&E systems and professionals are involved in the process.

The country and regional efforts that have been contributed by local M&E professionals have been impressive. These initiatives are directly addressing the challenge of a low enabling environment coupled with restricted M&E skills and trustworthiness, which places Arab countries in an opinion-based as opposed to an evidence-based system (Segone 2009). The region's M&E systems, capacities, demands, and professionalism are maturing. Dedicated local actors, such as the advocates and participants of EvalMENA, EREN, PEA, EvalJordan, RTE, LebEval, and EgDEval, are cultivating the culture for national evaluation processes to be aligned with other planning, budgeting, and statistics processes to drive the 2030 Agenda (El-Saddik et al. 2016). The demands of this era call for more international organizations and foreign donors to adopt the perspective of the IDRC and the Taqueem sponsors, in which local actors are regarded as partners with the capacity to support M&E and learning, and are appreciated as valuable assets to the development of their own countries and region.

¹⁰ Source: United Nations High Commission for Refugees website, <http://www.unhcr.org/en-us/syria-emergency.html>.

REFERENCES

- Abdallah, Abdulkhaleg. 2012. "Repercussions of the Arab Spring on GCC States." Research Paper, Arab Center for Research and Policy Studies, Doha, Qatar.
- Baradei, L., D. Abdelhamid, and N. Wally. 2014. "Institutionalising and Streamlining Development Monitoring and Evaluation in Post-Revolutionary Egypt: A Readiness Primer." *African Evaluation Journal* 2 (1).
- Boitnott, John. 2015. "How to Poach Talent Politely." *Entrepreneur*.
- Heliwell, John, Richard Layard, and Joffrey Sacha. 2015. *World Happiness Report 2015*. New York: The Earth Institute, Columbia University.
- El Kabbag, Nivine. 2011. "MENA Evaluation Experiences and Networks." Paper presented at IDEAS Global Assembly, April 2011.
- El-Saddik, Kassem, Dorothy Lucks, Stefano D'Errico, Thomas Schwandt, and Zenda Ofir. 2016. "Developing National Evaluation Capacities in the Sustainable Development Era: Four Key Challenges." IIED Briefing, International Institute for Environment and Development, London.
- ILO (International Labour Organization). 2009. "Making Learning Count: Effective Monitoring and Evaluation of Youth Employment Programs in the Arab World." Taqem Initiative. Geneva: ILO.
- Jansury, Laurel, Jennifer Moore, Juliana Peña, and Abigail Price. 2015. "Findings in Monitoring and Evaluations Practices During Humanitarian Emergencies." Paper prepared for International Business & Technical Consultants, Inc., Washington, DC.
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. "The Worldwide Governance Indicators: Methodology and Analytical Issues." Draft Policy Research Working Paper. World Bank, Washington, DC.
- Ministry of Foreign Affairs of Denmark. 2013. *Evaluation of Media Cooperation under the Danish Partnership Programme (2005–2012)*. Copenhagen.
- Moussa, Ziad. 2015. "Mainstreaming Evaluation Theory and Practice in the MENA Region." Final Report. International Development Research Centre.
- Office of Inspector General. 2015. "Survey of USAID's Arab Spring Challenges in Egypt, Tunisia, Libya, and Yemen." Survey Report No. 8-000-15-001-S. United States Agency for International Development (USAID), Frankfurt.
- Segone, Marco. 2009. "Enhancing Evidence-based Policy Making through Country-led M&E Systems. In *Country-led Monitoring and Evaluation Systems*, Marco Segone, ed. New York: UNICEF.
- Smith, Lydia. 2016. "Arab Spring 5 Years On: Timeline of the Major Events and Uprisings in the Middle East." *International Business Times* January 23.
- UN (United Nations). 2013. *The Arab Millennium Development Goals Report*. Beirut: United Nations Economic and Social Commission for West Asia (ESCWA).
- UNDP (United Nations Development Programme). 2015. *Human Development Report 2015: Work for Human Development*. New York: UNDP.
- UNDG (United Nations Development Group). 2011. *Results-Based Management Handbook*. UNDG.

Chapter 11

Leveraging Monitoring and Evaluation Systems for Good Governance

Rashmi Agrawal, Asela Kalugampitiya, Jigmi Rinzin,
and Kabir Hashim

Abstract. *Successful pursuit of the recently adopted Sustainable Development Goals by the world community depends on appropriate national development policies and actions. These should be guided by considerations of social equity, gender equality, and respect for environmental stability, and must be supported by good governance in order to contain corruption. This chapter discusses the concept of good governance, with an emphasis on the mutually dependent and mutually reinforcing relationship between good governance and sound monitoring and evaluation. Recent efforts and initiatives to bring about improved governance and monitoring and evaluation (M&E) effectiveness in three South Asian countries—Bhutan, India, and Sri Lanka—are discussed. Initiatives in India and Sri Lanka designed to contain corruption in public service delivery and make it more transparent, and the significance of M&E in promoting good*

Rashmi Agrawal, International Development Evaluation Association, rashmi_agrawal56@rediffmail.com; Asela Kalugampitiya, EvalPartners, asela.kalugampitiya@gmail.com; Jigmi Rinzin, National Council of Bhutan, jigmi@hotmail.com; Kabir Hashim, Sri Lankan Ministry of Public Enterprise Development.

governance and enhancing Gross National Happiness in Bhutan are described. These initiatives suggest some of the options available for leveraging M&E systems for good governance.

The adoption of the 2030 Agenda for Sustainable Development on September 25, 2015, at the United Nations Summit, which formalized the Sustainable Development Goals (SDGs) to be achieved by the world community of nations by 2030, has placed the SDGs center stage in the global development process. Attaining the 17 complex, interrelated SDGs in various socioeconomic sectors, and achieving the multitude of targets corresponding to each goal within the set time frame, has spurred action by the international community as well as by national governments to deliberate on what steps should be taken in order to make, measure, and assess progress toward these goals. While international action concerns sharpening of the goals and targets, and strengthening national capacities in measuring progress toward these goals, national governments are reviewing the capacities of their statistical and nonstatistical systems to respond to the requirements of such measurement. Attaining these goals calls for a systems approach, because the individual goals are not “in silos”: instead, they are mutually impacting, at times reinforcing, and on occasion, retarding other related goals. Multiple stakeholders—states, industries, civil society organizations (CSOs), and the community at large—must join hands and work together for this cause. While well-conceived policies and programs for equitable social and economic development through good governance can lead to sustainable development, sound monitoring and evaluation (M&E) systems would accelerate that progress. Policies and programs can lead to far better results under a scenario of good governance than when governance is poor. Good governance and effective M&E systems, which are mutually dependent and mutually strengthening, can contribute to optimizing returns on investment. While good governance assumes willingness on the part of policy makers and program planners and implementers to be subjected to a critical review of their actions, and to learn from such assessment, a robust M&E system can operate with a good measure of success only within an enabling environment. The imperative need for these two elements to work together has never been as obvious as it will be in the pursuit of SDGs by nations over the next decade and half. A robust M&E system that integrates within itself the requirements of sustainable development, mainstreamed within the national development agenda, is essential.

This chapter focuses on the concept of good governance and the interrelationships between good governance and sound M&E systems for sustainable development. It presents recent experiments and experiences in good governance and M&E in three South Asian countries—Bhutan, India, and Sri Lanka—and argues that there is more than one viable option available. The approaches followed in various countries can provide good learning that can be replicated in other places.

THE CONCEPT OF GOOD GOVERNANCE

What is “governance,” and what is implied by “good governance”? In Kautilya’s *Arthashastra*, good governance by a ruling king is described in these terms: “... in the happiness of his subjects lies his happiness, in their welfare his welfare.”¹

The United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) defines governance as “the process of decision-making and the process by which decisions are implemented (or not implemented)” in corporate, local, national or international contexts (ESCAP n.d.). Governance, therefore, is the result of the collective interaction of the decision making and implementing actions of the various actors and institutions concerned including those in both the public and private spheres of action: government at the national and local levels, industry, trade unions, CSOs, influential individual players, and even various organs of the media.

The World Bank, in its studies in more than 200 countries, has developed several indicators to measure the quality of governance using six dimensions: voice and accountability; political stability and absence of violence; government effectiveness; regulatory quality; rule of law; and control of corruption.² The World Bank views good governance as a necessary precondition for development, and the *Human Development Report* has defined good governance “as a democratic exigency [that], in order to [rid] societies of corruption, [gives] people the rights, the means, and the capacity to participate in the decisions that affect their lives and to hold their governments accountable for what they do” (Nzongola-Ntalaja, as quoted in UN DESA 2007, 4).

The United Nations has identified transparency, accountability, responsibility, participation, and responsiveness as the core attributes of good governance. Good governance is increasingly viewed as an essential element of any well-functioning society: when resources are allocated and used effectively, and the delivery of services to citizens in an equitable manner is ensured, the government gains a good measure of social legitimacy.³ The essential characteristics of good governance can be summarized as shown in figure 11.1.

Whatever the definition, good governance refers to a situation in which a set of institutions and actors combine to lead to sound processes of decision making, and the implementation of policies, programs, and projects that contribute to economic and social development, with “no one left behind.” “No one left behind” is the theme of the SDG agenda, and it is directly linked

¹ *Arthashastra*, a treatise in Sanskrit on governance and statecraft, is traditionally believed to have been authored by Kautilya (also known as Chanakya and Vishnugupta) c. 350–283 BCE.

² The Worldwide Governance Indicators project (<http://info.worldbank.org/governance/wgi/#home>) reports aggregate and individual governance indicators on these six dimensions of governance for more than 200 countries and territories over the period 1996–2016.

³ See United Nations Human Rights, <http://www.ohchr.org/EN/Issues/Development/GoodGovernance/Pages/GoodGovernanceIndex.aspx>.

FIGURE 11.1 **Components of good governance: a summary**

to good governance. The concept includes an element of flexibility in relation to its application in different country contexts: it may mean different things in different countries, depending on the cultures, traditions, political structures, economies, and levels of development. In a broad sense, good governance is an umbrella concept that covers respect for human rights, rule of law, an efficient and effective public sector, and processes of accountability and transparency of actions in public sphere.

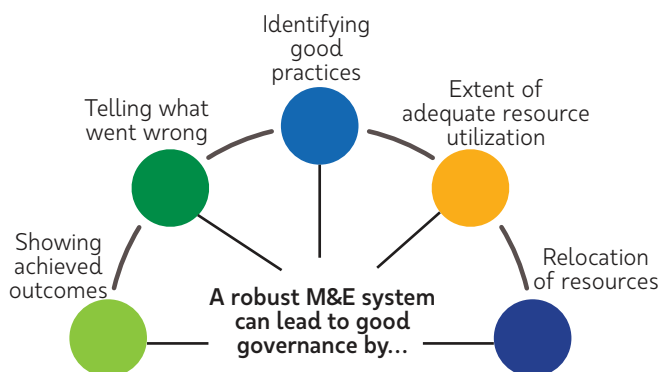
THE RELATIONSHIP BETWEEN GOOD GOVERNANCE AND M&E

There is a symbiotic relationship between good governance and M&E. Good governance creates an enabling environment for M&E, and M&E contributes to good governance. The former includes a keen desire on the part of the system to assess its own performance from time to time, to learn from experience, and to improve the outputs, outcomes, and impacts of the policies and programs it is pursuing. This is achieved through better planning, management, and implementation (PM&E): and this in turn creates a demand and an enabling environment for M&E.

Participation and responsiveness are two of the core attributes of good governance that also tend to increase the demand for M&E from civil society and other stakeholders. An enabling environment for M&E would also include the adoption of a clear national policy that promotes M&E in all its aspects, including the development of evaluation capacity, socially equitable

and gender-responsive elements, and the sustainability of evaluation. In turn, M&E works as a “reality check” tool for what authorities are saying about development, and what is really happening on the ground, and generates evidence-based lessons for the future, thereby contributing to knowledge, and suggesting policy and program modifications for enhanced future outcomes. Good governance provides a way for good M&E systems to exist and evolve, while M&E systems provide governments with evidence and learning that helps in need-based policy planning, and the improvement of ongoing programs and learning. The relationship of good governance and M&E can be better understood from figure 11.2.

FIGURE 11.2 **Relationship between good governance and M&E**



GOVERNANCE AND M&E IN INDIA: AN OVERVIEW

India’s constitution provides for a republic with a democratic, secular, and socialistic form of society: it places the principles of universal equality and social justice on a high pedestal. Appropriate affirmative action by the administration aims to raise the standard of living for the less socially and economically advantaged so that the gap between them and the rest of the society tapers off, and they can join the mainstream of development. The government aims to translate this intention into reality through a series of social and economic development policies, plans, and programs, through massive investment in the social sector. But huge investment does not necessarily produce the desired outcomes, unless it is operated in a scenario of sound governance. Good governance is critical to ensure that these investments lead to significant outcomes and impacts on the ground, through the efficient use of allocated resources; optimal management of public service delivery; and effective management of natural resources.

In a democracy, rising income levels also bring with them rising expectations among citizens, and a demand for good governance at all levels of the government: national, state, and local. Experience over the past over six decades indicates that while significant progress has been made in India

on most of the economic and social parameters, the impacts are not commensurate with the resources utilized, and could have been vastly superior with better governance. A significant step-up is required, through systemic improvements in implementation; increased efficiency of public agencies in the delivery of services to consumers; and tackling the menace of corruption, which has siphoned off huge chunks of public investments.

The Indian system of governance has two types of actors—formal and informal. **Formal actors** include the national and state legislative bodies, the judiciary, government functionaries, and constitutional bodies such as the Election Commission, the Comptroller and Auditor General, the Central Vigilance Commissioner, and organized industries and services in both the public and private sectors. **Informal actors** are the multitude of civil society organizations, academicians, the media, and the community. In the democratic set-up in India, both of these types of actors play an important role in planning and implementation. Together they contribute to the governance of the land as policy makers, enforcers of the enacted policies and laws, program planners, providers and receivers of various goods and services, and guardians of citizens' rights. In terms of economic and social development, there is broad consensus concerning the goals of inclusive and sustainable growth and gender equality, in line with global objectives and standards.

The National Institution for Transforming India (NITI Aayog), an organization that replaced the 60-year-old Planning Commission, is currently the main think tank for Indian policy makers on developmental issues.⁴ NITI is looking forward to maintaining a state-of-the-art resource center, to be a repository of research on good governance and best practices in sustainable and equitable development, as well as to help them disseminate results of such research to stakeholders. NITI is also responsible for actively monitoring and evaluating the implementation of programs, and for identifying the resources needed to strengthen the probability of success and the scope of delivery. This is the main arm for M&E in the country: thus it is aimed at both strategic policy and program frameworks, and is monitoring both their progress and their efficacy.

Recent Initiatives for Good Governance in India

India has a long history of policies and programs directed toward economic and social development, tuned to the principles of inclusiveness (five-year plans; for instance, see Government of India 2013). However, despite lofty ideals and goals, performance has often fallen short of expectations. The present government came to power with a motto of “minimum government and maximum governance.” In its objective of providing good governance, it has been seeking to simplify the delivery of services to citizens, and make the process as transparent and corruption-free as possible, primarily through the application of information technology as an interface between the

⁴<http://niti.gov.in/content/overview>.

government and citizens. Some recent initiatives aimed at these goals can provide learning.

Digital India. Digital India is a flagship program of the government of India, with a vision of transforming India into a digitally empowered society and knowledge economy.

The program builds on various e-governance initiatives for wider sectoral applications, with an emphasis on citizen-centric services. The main pillars of the program include universal access to mobile connectivity, public Internet access, e-governance, electronic delivery of services (*e-Kranti*), and information for all.

Jan Dhan Yojana.⁵ A vast majority of the Indian population have traditionally depended on informal financial services and remained outside of the formal banking system, making it difficult for service providers to reach them through formal channels. This informal system also gave birth to corruption. Jan Dhan Yojana is an important step taken by the government to mainstream that part of the population that has previously not been covered by banking services into the formal system by enabling and encouraging them to do so through incentives. Regular monitoring has indicated that until December 21, 2016, about 260.3 million savings bank accounts were opened under the scheme, 158.6 million of them in rural areas, and 101.7 million in urban areas, accounting for a total deposit balance of Rs 7,155.7 billion (or about \$110 billion). This is a significant step toward financial inclusion that will facilitate seamless and direct transfer of subsidies and other benefits into beneficiaries' accounts, reducing the number of opportunities for funds being pilfered en route.

De-monetization and Digi Dhan Yojana. Over 80 percent of the country's economic activities are carried out in the informal sector, through informal financial and other transactions, with no accounts kept. This state of affairs leaves enormous room for underreporting or nonreporting of incomes, thus undercutting the potential for tax collections; allows for corrupt practices of paying in cash for irregular services to take place; and enables some people to accumulate enormous amounts of black money.⁶ This money, in large quantities and in fake currency, enables the funding of drug-related activities, terrorism, and other antisocial activities. A recent action by the government aimed at dealing with corruption, black money, and tax evasion has been de-monetizing high-denomination currency of India (notes of value Rs 1,000

⁵ Jan Dhan Yojana (literally meaning "people's money program") is aimed at ensuring access to various financial services including availability of basic savings bank accounts, access to need-based credit, remittances facility, insurance, and pension to excluded populations (e.g., vulnerable and low-income groups).

⁶ Black money points to incomes and wealth from undisclosed and often illegal transactions, and on which taxes are not paid.

and Rs 500), which together accounted for about 85 percent of all cash in circulation, and issuing fresh currency to replace the deposits of old currency in the banks. Though the process has resulted in some temporary inconvenience to the people, long-term gains are expected, through increased tax revenue and a reduction in corrupt practices. There has been a visible decline in terrorist activities.⁷

Simultaneously, the government has launched a massive program to enable and encourage both consumers and traders to learn and to use noncash (digital) means of money transfers. It is expected that this change-over to an economy that is less dependent on cash transactions would make transactions more transparent, boost tax income, and make it easier to curb corrupt practices. A continuous monitoring process has helped to assess the problems coming out of these new initiatives, and to take corrective actions immediately. For example, to ensure that the lack of Internet connectivity and education do not hamper the practice of digital money transactions, a new app, Bharat Interface for Money (BHIM), which does not require Internet connectivity and can be used even by people without education, has been launched.⁸

Clean India Mission. This is another major initiative of the government aimed at making India open defecation-free by 2019 and making people aware of the importance of keeping their environs neat and clean, through an aggressive awareness-generation campaign involving prominent people from all walks of life, and providing financial assistance for building household and institutional sanitation facilities. M&E is a regular part of this program. Cities, towns, and villages are being ranked according to their level of cleanliness. Success stories are being disseminated across the country, and the work is going on with active community participation.

Make-in-India and Skill India. Other major developmental initiatives with an equity focus include the Skill India and the Make-in-India programs. The main goal of Skill India is to create opportunities, space, and scope for the development of the talents of Indian youth, and to further develop those sectors that have already been playing a role in skill development for the last so many years, and also to identify new sectors for skill development. This new program aims to provide training and skill development for 500 million youth by 2020, and to cover each and every village.

⁷ See, for instance, *India News* October 8, 2017, referring to the finance minister's observations on the subject, www.dnaindia.com.

⁸ The BHIM app enables people to make simple, easy, and quick payment transactions using a unified payments interface (UPI) with just a mobile number or UPI ID. Pioneered and developed by the National Payments Corporation of India, BHIM was launched by the prime minister, Narendra Modi, on December 30, 2016, to usher in a financially inclusive nation and a digitally empowered society.

The purpose of Make-in-India is to encourage local and international manufacturers to set up production facilities within the country to boost production and employment.

All of these initiatives include M&E components: thus, M&E is being mainstreamed into the development process.

Strengths and Weaknesses of M&E in India

India, with its history of more than six decades of developmental planning supported by a network of M&E institutions and activities, is not new to this sphere of activity. But the growing importance of SDG-oriented M&E demands a stronger and more focused approach to M&E, supported by an M&E-enabled atmosphere, capacity building, and evaluation knowledge-sharing mechanisms.

Institutional mechanisms. The Development Monitoring and Evaluation Organization, a component of NITI Aayog, aided by its 15 regional offices across the country, keeps track of the developmental agenda in the context of SDGs, with evaluation as a priority. Apart from this centralized institution, most of the ministries and departments in the national and state governments have their own M&E systems.

Capacity building in M&E. Multipronged efforts are being made to develop national M&E capacities. A number of national institutions, such as the National Institute of Labour Economics Research and Development (NILERD), and international institutions such as the International Initiative for Impact Evaluation (3ie) and the Centers for Learning on Evaluation and Results (CLEAR) are organizing workshops, sensitization programs, and both short and long-term training programs for this purpose. Various states have approached NILERD, asking them to organize short-term programs for their officials. This indicates a growing awareness of the need for M&E capacities.

Toward a national evaluation policy. In spite of a long history of development evaluation, and the amount of emphasis currently being placed on M&E, India still does not have an explicit national evaluation policy. The strong need for a national evaluation policy that will provide a framework that defines the principles governing the role of M&E in development; the approach, quality, methods, and ethics to be ensured in the practice of development evaluation; utilization of the evidence-based results of such evaluations; and, importantly, the human and material resources to be optimally allocated for this purpose, is increasingly being recognized.

The Evaluation Community of India. Due to the growing demand for evaluations, and for capacity building from various corners, it was felt there was a need for a platform where planners, implementers, evaluators, and communities could come together to discuss various issues relating to M&E. Such a platform was launched in 2015, through the Evaluation Community of India (ECOI), a voluntary organization for professional evaluation (VOPE) with

the motto “to share and learn.” ECOI has action groups working on various aspects of evaluation, such as preparing a draft national evaluation policy, capacity development in evaluation, and so on. ECOI is looking forward to networking and interactions with various partners to collaborate in further developing evaluation culture.

RECENT DEVELOPMENTS IN M&E AND GOVERNANCE IN SRI LANKA

Over the years the important role of M&E has been well recognized by the government of Sri Lanka.⁹ Sri Lanka was one of the first countries to establish a web-based project monitoring system, e-PMS, to track projects implemented across all ministries. A homegrown electronic system was a significant aspect of that set-up. It was established in the then Ministry of Plan Implementation, to track financial and physical progress in implementation, and the results of all development projects and programs. The system could generate project information donor-wise, sector-wise, and ministry-wise. The Department of Project Management and Monitoring (DPMM), which has the mandate for M&E, has now replaced this system with a new Integrated National Development Information System.

However, the system is more or less confined to output-based progress monitoring of various development programs. In the context of SDGs, and the growing demand for effective monitoring and higher-level impact evaluations, the need for a comprehensive national M&E system is being increasingly felt. The Global Evaluation Agenda that evolved at Kathmandu in 2015 (Eva-IAgenda 2020) aims to strengthen the enabling environment for evaluation; develop institutional capacities; build individual capacities for evaluation; and support links among these first three dimensions by all stakeholders—governments, parliamentarians, VOPEs, the United Nations, foundations, civil society, the private sector, academia, and other interested groups—working together. In response to these developments, the Parliamentarians’ Forum for Development Evaluation (PFDE) South Asia project team organized a series of events to facilitate the establishment of a national M&E system in Sri Lanka. Groundwork for such a system has recently been laid by the initiatives of VOPEs, the Sri Lanka Evaluation Association, the South Africa Monitoring and Evaluation Association, and the Malaysia Evaluation Association, together with their respective government agencies, and with the support of the Eval-Partners Peer-to-Peer Small Grants Programme.

Sri Lanka’s National Evaluation Policy

Although Sri Lanka commenced evaluations in the 1990s, the National Evaluation Policy (NEP) process had not been continued consistently until recently: one reason for the delay was the lack of its endorsement by the government. The DPMM is the department within the Ministry of National Policies and

⁹Paper submitted by Priroshini Trikawalagoda to the APEA conference 2016.

Economic Affairs that has the mandate for M&E, under the leadership of the Prime Minister. DPMM was stimulated to take a lead in the NEP process after the implementation of EvalPartners Peer-to-Peer Small Grants Programme 2015: one of outputs of this was the draft preliminary action plan formulated to roll out the NEP. A draft of the policy initially developed by Sri Lanka Evaluation Association is being refined through a series of consultative processes between DPMM and other stakeholders that has been organized by PFDE and supported by EvalPartners, the EvalGender+ initiative, and UNICEF. A road map—another output from the stakeholder consultation—is expected to help guide the process for obtaining Cabinet approval for the NEP.

Motions in the Sri Lankan Parliament

Two significant and highly encouraging events have recently taken place in the Sri Lankan Parliament. Two adjournment motions were moved in the Parliament: one to formulate a NEP, and the other to allocate funds for evaluation. The first motion was made in August 2016 by a member of parliament who proposed formulating a NEP, and an evaluation system for the country to assess whether the anticipated results from development programs have been achieved. Making the motion, the honorable member stated that Sri Lanka had pledged to achieve the SDGs by 2030, and emphasized the importance of the role of evaluation in assessing whether the anticipated results from development interventions have been achieved. He cited examples from several countries where evaluation systems have been established, and stated that around 20 countries have already formulated NEPs, demonstrating his interest in and commitment to the cause. The same member also called for a separate allocation of funds from the national budget to be set aside for evaluation. All of the parliamentarians who participated in the debate were in favor of both motions. It seems likely, therefore, that the Sri Lankan Parliament is keen to adopt a NEP.

The parliamentarians who joined the debate agreed in unison that the present national government, with the president and the prime minister representing the two leading parties, presents a good opportunity for implementing a NEP. PFDE–South Asia is closely supporting reactivation of the NEP process, to be backed by evaluation capacity development. Being the heads of the ministries concerned with the subject, the prime minister and his deputy are encouraging the initiatives.

During the debate, Malaysia, Nepal, and South Africa were cited as examples of countries that had commenced the process ahead of Sri Lanka. It is noteworthy that globally this may be the first-ever motion on evaluation moved in a parliament. Nevertheless, if the evaluation process is to function independently, a systematic framework that includes ministries, departments, boards, and so on, would have to be developed. For this purpose, a sizable allocation of funds is needed. These funds are worthy investments, since evaluation will facilitate the achievement of the objectives of the projects. Furthermore, the lessons learned from the process will help in effective decision making in the short, medium, and long term for the projects being evaluated, as well as when implementing future projects of a similar nature.

An Online Web-Based Project Monitoring System

During the stakeholder consultations, the need to integrate areas of interrelated work, evaluation, data, and indicators with a strong information system that would be backed by relevant, high-quality, disaggregated, comparable, and timely data was emphasized. The existing system needs to be modified and upgraded so that it can cater to current needs. However, because it has no local partner the cost of maintenance is exorbitant. The system has been handed over to the Information Communication and Technology Agency, the government's information and communications technology arm, which is in the process of entering into a memorandum of understanding to modify and upgrade to a more user-friendly system. Some of the other issues currently faced by DPMM are inadequately trained staff in line ministries due to transfers to other departments with no succession planning, and delays in updating the system by the line ministries. Furthermore, it is difficult to verify the accuracy of data when the observations of senior management are not entered into the system. More stakeholders would be encouraged to use the system if it could be made more user-friendly, and if project progress could be tracked and used to contribute to informed decision making. PFDE–South Asia has also recognized the need to build awareness and strengthen the Department of Census and Statistics, and has invited them to participate in evaluation capacity development programs.

Capacity Building of Public Sector Officials

In Sri Lanka, several capacity-building initiatives have taken place, including training workshops for government officials, and study tours to Malaysia and South Africa. The training workshops included evaluation, the Logical Framework Approach in evaluation and designing and managing evaluations.

A four-member delegation comprised of representatives of DPMM, and two representatives from PFDE–South Asia, visited South Africa with the objective of developing a preliminary action plan to implement a national evaluation system in Sri Lanka. The delegation had the opportunity to learn about the South African system of government and the application of planning, monitoring, and evaluation by the South African Department of Performance Monitoring and Evaluation. Discussions included the timelines used in developing a macro PM&E system, the rationale for PM&E, the various roles and responsibilities of different organs of the government, the outline of the department's work in the context of the central government, the National Development Plan 2030, the National Evaluation Plan, and the problems PM&E aims to address. The web-based system of frontline monitoring of the presidential hotline for community problem solving was also demonstrated.

The delegation also visited a participatory workshop on preparing terms of reference for an integrated development plan to ensure safety in the Republic of South Africa where they learned about the Management Performance Monitoring Tool (MPAT) and how performance is measured in departments using scorecards. The study tour was an important step in the long-term evaluation capacity-building program in Sri Lanka, supported by

the EvalGender+ network and the UNICEF country office. Such international experiences are very important if countries are going to work together to achieve the SDGs within the scheduled time frame.

In short, substantial efforts have been made to develop an evaluation culture in Sri Lanka, but there is a still need for the following:

- Strengthen parliamentarians' desire to use and demand evaluation
- Strengthen the evaluation capacity of district development committees where a decentralized budget is used
- Government endorsement of the national evaluation policy

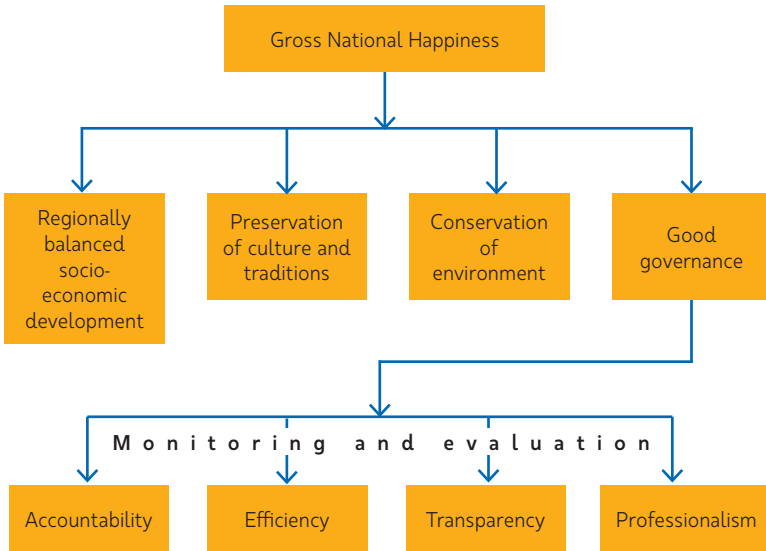
GOOD GOVERNANCE AND M&E IN BHUTAN: OVERVIEW

Bhutan's economic plans and programs are guided by an overarching development philosophy of Gross National Happiness (GNH). The concept of GNH attempts to ensure that an economic activity not only reaps material benefits but also positively impacts issues like equity, sustainability, preservation of the environment, and culture. GNH is supported with four pillars: good governance is the fourth pillar, which is the underpinning for the success of the other three pillars.

Good governance in Bhutan is characterized by four features: accountability, transparency, efficiency, and professionalism. While the concept of M&E is a cross-cutting phenomenon transcending all four pillars of GNH, the practice of M&E is more obvious and apparent with respect to good governance. Evidence-based practice of good governance is desirable for ensuring success of the other three pillars. Figure 11.3 presents a flowchart depiction of GNH, good governance, and M&E in the Bhutanese context.

Because of its success in achieving the Millennium Development Goals, under the development strategies of the GNH philosophy, Bhutan involuntarily has already begun the implementation of development activities for the SDGs. This is due to the fact that under the GNH development strategies, activities desired under the SDGs have inherent links with existing development plans and programs. For instance, the GNH pillar concerning "regionally balanced equitable socioeconomic development" has inherent links with SDGs 1, 2, 3, 4, 5, 7, 8, 9, 10, and 11. Similarly, the pillar of "preservation and promotion of culture and tradition" has links with SDGs 11 and 12. The third pillar, "conservation of the environment" can conveniently absorb SDGs 6, 7, 12, 13, 14, and 15, while the "good governance" pillar is related to SDGs 16 and 17.

Bhutan's Planning Commission, known as the GNH Commission (GNHC), is the central coordinating agency for development PM&E. In a bid to streamline and institute an effective system for M&E of development plans in the country, in 2006 the GNH Commission developed a national M&E system, as a standard system for monitoring and evaluating the development plans and programs administered by ministries and agencies. A dedicated unit, known as the Research and Evaluation Division, is in place at the GNH Commission.

FIGURE 11.3 **GNH, good governance, and M&E in Bhutan**

The Status of M&E in Bhutan

A strengths, weaknesses, opportunities, and threats (SWOT) analysis of Bhutan's M&E system was undertaken during one seminar on evaluation in Bhutan in March 2013, conducted by the GNHC and UNICEF-Bhutan. It was found that the evaluation system was weak; the technical capacity to conduct, commission, and manage evaluations was lacking; and the demand for evaluation was low. It was also noted that evaluations in Bhutan were generally donor-driven. These factors posed challenges for strengthening the evaluation culture in Bhutan. It was also recommended that a nonprofit association and a network of evaluators be established, to provide the much-needed platform to promote evaluation in Bhutan.

Therefore, evaluation is still at the nascent stage to this day. Evaluation is less understood and appreciated. Neither is there a demand for evaluation nor the supply. To this extent, there is a lack of capacity in evaluation.

On the contrary, there has been a considerable progress on the monitoring aspect. The sense of monitoring is not only becoming firmer in the system, but it is easier and more straightforward for agencies for implementation and oversight purposes.

In order to streamline, strengthen, and institutionalize the evaluation system, the national evaluation policy, and the national evaluation protocol and guidelines have already been formulated, and are awaiting the government's approval. Several evaluations of development policies and programs have also been conducted since 2013, through the Research and Evaluation Division of the GNHC, and in collaboration with government ministries. These

initiatives are geared toward promotion of the demand and use of evaluation by governments and parliaments so as to inform policy development and increased social accountability to citizens through evaluation.

Bhutan's Planning and Monitoring Process

Bhutan's development plans and programs are based on overall five-year plans (FYPs); currently, the country is in its 11th FYP (2013–18). FYPs are prepared through government consultations with implementation agencies, both at the national and grassroots levels. Plan consultations are preceded by issuance of planning guidelines to agencies. As part of results-based planning and management practices, plans are corroborated with identified result areas known as national key result areas, sectoral key result areas, key performance indicators, and specific key interventions. Individual FYPs are prepared by the respective ministries, autonomous agencies, and local government agencies, and their plans must be aligned with the national priorities. Based on the approved FYPs, annual work plans (AWPs) and budgeting are prepared and executed during the year.

Appropriate information technology (IT) systems are employed at various stages in the planning and monitoring continuum. The formulation of overall FYPs is based on information provided by planning and monitoring systems (PLaMS). PLaMS also provide support during the preparation of AWPs and their implementation.

Annual budgets are prepared using a multiyear rolling budget system. Budget releases are made on a quarterly basis, upon receiving plan monitoring and progress reports, which are essentially both physical and financial progress reports. While these reports will be made by implementing agencies through the PLaMs, the Ministry of Finance will release periodic budgets through another IT system known as the public expenditure management system (PEMS). Every government transaction is conducted online through PEMS, upon submission of progress reports. Thus, the AWPs and the budget, during the implementation phase only, will have periodic monitoring and reporting requirements that must be followed by the agencies.

It is customary for the government to conduct midterm reviews of FYPs in the middle of the Five-Year period, typically when the plan has progressed two and a half years into the FYP. In addition, the present government has initiated the drawing of annual performance agreements between the prime minister and individual agencies to ensure successful implementation of the AWPs. The activities identified in the AWPs would be determined by the respective agencies and duly agreed upon with the prime minister. There will be an annual review of annual performance agreements with the implementing agencies.

Status of the Evaluation Profession in Bhutan

Reaffirming the importance of evaluation within the system, UNICEF-Bhutan has been engaged along with the GNH Commission since the first day of the dialogue on promoting evaluation in Bhutan. Equally, parliament, through its

various standing committees, has always been on the frontline to promote evaluation culture in the country.

In an attempt to promote evaluation culture in Bhutan, with the technical and financial support of UNICEF-Bhutan, and the administrative support of the GNH Commission, the Evaluation Association of Bhutan (EAB) was formed in 2013, with a multibackground membership. Its registration for formal recognition of CSO status is already in process. In collaboration with UNICEF, the Community of Evaluators-South Asia, and the GNH Commission, the EAB regularly conducts training and experience-sharing events for its members and other stakeholders. Thus, the EAB is working on creating a network of high-quality evaluators and linking them with other evaluation communities.

From 2009 until mid-2015, about 16 officials from the government had attended the International Program for Development Evaluation Training (IPDET) funded by various sources of training, a major portion of which was supported by Danida. This training contributed immensely to enhancing awareness and capacity in the Royal Government of Bhutan. About 24 Bhutanese with an interest in evaluation are members of the International Organization for Collaborative Outcome Management, which was established in 2010. In 2013, a few Bhutanese evaluators joined the Community of Evaluators as individual members: this has strengthened the discourse on development evaluation, and the evaluation culture in the country.

However, despite these steps forward, the evaluation profession has not progressed very much. Except for donor-funded programs, no evaluation by independent practitioners has been conducted for government programs. And even for donor-led evaluations of programs and projects, the evaluations are mostly carried out without established processes and standards. Nonetheless, there is a silver lining, with the government's relentless effort toward the promotion of both demand and use of evaluation, and the disclosing of policy development through increased social accountability mechanisms.

The more plausible solutions, at this stage, are to:

- Expedite formal government approval of the national evaluation policy, guidelines, and protocol;
- Recognize the EAB as a legal CSO;
- Upgrade the capacity of evaluation practitioners; and
- Enhance the utilization of evaluation reports by agencies.

CONCLUSION

Bhutan, India, and Sri Lanka are all South Asian countries, but their approaches to governance and M&E vary. Analysis indicates that Sri Lanka is still confined to an output-based monitoring system, although recent events suggest a growing realization of the importance of, and gradual transition to, higher levels of evaluation. The emphasis in India, with its long-entrenched M&E systems, has been on outcome monitoring and impact evaluations. However, it has no evaluation policy. This often results in inadequacy or a multiplicity of efforts in M&E, a lack of standardized practices of evaluations, and more

importantly, their inadequate utilization. Sri Lanka on the other hand, has initiated the process of developing an NEP, and outcomes have found a place in the national parliament. Sri Lanka has adopted a long and robust process of dialogue and consultation with various stakeholders to make their evaluation policy a reality, and a motion for specific fund allotment for evaluation has been raised in their parliament. In India, there is a strong need for evaluation policy, and the allotment of dedicated funds for evaluations to make them utility-oriented, as highlighted in recent stakeholder discussions. In recent times, India has witnessed initiatives toward good governance and promotion of the use of IT to make citizen services efficient and financial transactions transparent, thereby reducing corruption. Bhutan, meanwhile, has a totally different philosophy of good governance, and measures it by public happiness. Bhutan considers the sole purpose of development is making people happy. High incomes may lead to material benefits, but general happiness is more important than these material benefits. Therefore, in Bhutan M&E is a cross-cutting issue looking to assess impacts in terms of GNH.

All three of these countries are committed to achieving the SDG targets; and M&E, as well as good governance are essential tools in that process. All three countries share the view that capacity building in M&E is extremely important in order to achieve the SDGs. While in India evaluations are getting mainstreamed into the development agenda and there is a demand for evaluation from various stakeholders, in Sri Lanka and Bhutan evaluations are generally donor-driven. Now is the time for various countries to come together, share their experiences, and learn from each other for future action.

REFERENCES

- ESCAP (United Nations Economic and Social Commission for Asia and the Pacific). n.d. "What Is Good Governance?" Fact sheet. ESCAP, Bangkok.
- Government of India. 2013. "Twelfth Five-Year Plan (2012–2017), Vol. I: Faster, More Inclusive and Sustainable Growth." New Delhi.
- UN DESA (United Nations Department of Economic and Social Affairs). 2007. "Public Governance Indicators: A Literature Review." ST/ESA/PAD/SER.E/100. United Nations, New York.

Chapter 12

Integrating Feminist Approaches to Evaluation - Lessons Learned from an Indian Experience

Rajib Nandi and Rituu B. Nanda

Abstract. *Feminist approaches to evaluation seek to unpack the nature of gender and social inequalities; treat evaluation as a political activity, not as a value-free assessment; and use it as part of the change process. In response to the increased attention being given to evaluation, and the possible role of feminist evaluation in influencing policy that could lead to gender equality along different dimensions, the Institute of Social Studies Trust, based in New Delhi, designed a four-year program. The aim of program was to enhance capacity and understanding of feminist evaluations of various stakeholders, with a conviction that it would influence gender-transformative policy making. This chapter discusses the objectives and activities of this program, and critically analyzes the lessons learned. It also highlights the important takeaways that can make evaluations transformative as far as gender-based inequalities and power dynamics are concerned.*

Rajib Nandi, Institute of Social Studies Trust, nandi.rajib@gmail.com; Rituu B. Nanda, Institute of Social Studies Trust and Constellation, rituubnanda@feministevaluation.org.

The Indian economy and society have been struggling with persistent gender gaps and inequalities. Despite substantial progress over the years, gender gaps continue to persist in education, health, participation in the workforce, and decision making (Hay et al. 2012). Poverty, early marriage, malnutrition, and lack of health care during pregnancy are associated with high levels of maternal and infant mortality. Data show that in India almost 60 percent of girls are married before the age of 18 (UNICEF 2014), and nearly 60 percent of them bear children before they are 19 (Young Lives 2016). In addition, almost one-third of all babies are born with low birth weight. Although gender parity in school enrollment has largely been achieved, there are gender differences in the reasons for dropping out, for irregular school attendance, and for the pathways that open up through education. Against a male workforce participation rate of 53.0 percent in rural areas and 53.8 percent in urban areas, the female workforce participation rate was 30.0 percent and 15.4 percent in 2011 (GOI 2016). Women continue to be employed mainly as “marginal” workers, in home-based, informal economy work, and as unpaid family labor. They remain underrepresented in decision-making positions, even though quotas have enabled more than a million women to enter local governance institutions.

Because gender norms affect all aspects of work and life, and gender is seen as a cross-cutting issue, it often becomes invisible in policy and planning documents. For example, neither India’s midterm appraisal of the Eleventh Plan, nor the issues for approach to the Twelfth Plan, mention “gender equality and empowerment” as a separate tangible goal, and discussion with planners reveals that it is seen as a “cross-cutting” factor (Planning Commission 2011b, 2011c). The reluctance to make gender concerns more explicit, which is derived from ignorance of the facts about gender gaps, suggests that there is still significant doubt as to whether any alternative policy or program design would significantly alter gender-related outcomes or not. However, evidence from the work of feminist and gender-sensitive social scientists shows that identifying and addressing the factors leading to gender inequalities can indeed help develop strategies for both policy advocacy and implementation for better outcomes from a gender perspective.

The present chapter highlights the discussions and lessons learned from a four-year program of capacity building on feminist evaluation that was prompted by an apprehension that evaluations with a feminist lens are capable of offering evidence-based policy advocacy that is oriented to gender equity and social justice.

The authors consulted different program documents including the program proposal, narrative reports, workshop reports, and the evaluation reports drafted by the external evaluators in order to write this chapter.

ENGENDERING POLICY THROUGH EVALUATION: BACKGROUND AND OBJECTIVES

Gender in Indian Policy and Planning

The ways in which a gender and equity lens in evaluation can help improve policies and programs is an important message for policy makers to receive.

On the other hand, it is widely recognized that evaluation has been emerging as a critical space through which gender and equity questions can be brought back into policy discourse.¹ The Indian government's Eleventh Plan adopted a gendered lens to initiate a process of systemic improvement in the lives of women and children. But the plan's midterm appraisal shows that while certain sectors have shown remarkable improvement, others are lagging behind (Planning Commission 2011c). The approach to the Twelfth Plan shows that only 35.8 percent of the Eleventh Plan outlay has been allocated during the first three years of the plan (Planning Commission 2011b). Schemes for single and internally displaced women, domestic workers, and minority women, to name a few, have not found a voice in the first half of the Eleventh Plan (Planning Commission 2011a).

During the formulation process of the Twelfth Plan, the coalition of civil society organizations and the planning commission was strengthened. The planning commission invited civil society groups to contribute to the preparation of the approach paper for the Twelfth Five-Year Plan, which is aimed at achieving faster, more sustainable, and more inclusive growth. The United Nations Development Programme (UNDP) India supported a facilitation process that provided women and men from marginalized communities living in remote corners of the country with an opportunity to voice their opinions on key development issues, and in doing so, marked a significant step in making the planning participatory. Planners interacted with 16 population groups comprised of *dalits*,² migrants, the urban poor, ethnic minorities, Muslims, people living with HIV/AIDS, and transgender persons, among others (WNTA 2011).

Consequently, the Twelfth Plan approach paper acknowledged that the plan must break the vicious cycle of multiple deprivations faced by girls and women because of gender discrimination and undernutrition. The paper ensures that ending gender-based inequities, discrimination, and violence faced by girls and women must be accorded the highest priority, and that this needs to be done in several ways. The midterm appraisal report of the Eleventh Plan acknowledges a need to undertake impact evaluations of intended outcomes. To undertake such evaluation, it has been decided to establish an independent evaluation organization linked to, but distinct from, the Planning Commission (Planning Commission 2011c).

Within this environment, there is some evidence that greater attention is now being paid to government-commissioned independent evaluations, as well as to the methods used and the findings. Acknowledgment by the government has created a tacit space for discussing gender-responsive evaluations in India. A meeting of the evaluation agencies and implementing agencies of the Support to Training and Employment Programme (STEP) for Women of the Indian government's Ministry of Women and Child Development was

¹ *Transform: The Magazine for Gender-responsive Evaluation*, Issue 1, June 2015. UN Women, Independent Evaluation Office.

² *Dalit*, meaning "oppressed," is a term for the members of lower castes of India.

convened in May 2011 for the first time since the program had started in 1986, to reflect on the design and evaluation processes of the program. The participants provided their thoughts about a systemic development of the evaluation process for the program. The increasing demand for greater transparency and openness is further reflected in the recent citizen agitation demanding an anti-corruption bill,³ and in the use of the Right to Information Act to ensure public accountability. The act is becoming more crucial for the vulnerable sections of our society as they battle for social and economic justice. This is particularly true for women's issues. Some recent reports show that groups of women have also been able to fight gender-based discriminations through the Right to Information Act (Bakshi and Bhattacharya 2010).

A meta-evaluation of the STEP was conducted in 2012 by the Institute of Social Studies Trust (ISST), using a feminist lens. The study points out that care responsibilities and constraints on mobility can influence outcomes, and need to be factored into assessments of both design and outcomes. This study was an attempt at a formative meta-evaluation using a synthesis method, and with the purpose of using completed evaluation reports to inform and strengthen future evaluations ("Introduction," Sudarshan, Murthy, and Chigateri 2015).

The meta-evaluation of the STEP has provided critical insights, as follows:

...the importance of assessing whether the trainings were contextualized to the sector and catered specifically to the needs of the women was also emphasized. The argument made by the evaluating agency was that the trainings would not be effective otherwise, and that women would just remain recipients of STEP. In order to empower the women, the trainings had to be linked to the lives of the women. Furthermore, the trainings had to cater to the functional requirements of the women. For instance, it was recommended that the legal training not be limited to awareness of legal rights but should also address functional and transactional legal literacy in the particular context of the sector, for instance to know what happens in the case of non repayment of loans, or to any other specific issue from a particular sector like farming or weaving etc. Therefore, in order to assess the effectiveness of the training component, it is also important for the evaluating agency to assess whether the training was contextualized and catered specifically to the needs of women. (Chigateri et al. 2015, 65)

Importance of Building Evaluation Capacity in India

There is a strong presence of international evaluators, both organizations and individuals, in India today. This has helped to generate a more visible discourse about evaluation. Questions are being raised regarding things such

³ See, e.g., news reports at <http://www.thehindu.com/news/national/article1607789.ece>.

as the extent to which greater professionalization of the evaluation function is needed, and how important the role of contextual understanding and domain knowledge is.⁴

Whether there is a growing demand for the evaluation of programs and projects, or whether there is simply greater visibility of this area of work, there is a sense among many commissioners of evaluations that local evaluation capacity is weak: this means that even if more evaluations are being locally commissioned, there may be a sense of discomfort among the commissioners regarding their credibility. Shiva Kumar, for example, has commented that “Professionals carrying out evaluations in South Asia tend to be social science researchers, not trained evaluators. Many evaluators of development interventions and commissioners of evaluation have not fully realized that the competencies needed to become an evaluator are different from, though complementary to, those needed for conducting social science research” (Shiva Kumar 2010).

Gender-transformative approaches to evaluation seek to unpack the nature of gender and social inequalities: further, they see evaluation as a political activity, not as a value-free assessment, and use it as part of the change process. Responding to the increased attention being given to evaluation, and the possible role of feminist evaluation in engendering policy and supporting changes that lead to gender equality along different dimensions, a consultation on gender and participatory evaluation was organized by ISST in August 2010 in New Delhi. There is, so far, little writing on gender-transformative evaluation tools and frameworks, and little on the difference made to evaluation findings by using a feminist lens. This workshop brought together a group of feminist and gender-sensitive equality advocates who, as social science researchers, have carried out evaluations and not had the opportunity to reflect on the role of these evaluations within their larger research agendas. Examples were given of the ways in which evaluation has been able to shift policy perspectives through the redesign of programs, and more importantly through systems change.

The participants of the above consultation might have lacked familiarity with mainstream evaluation theories and theorists, and the tools and frameworks associated with them, but the presentations by evaluators at the workshop showed that they have a good knowledge of the concerned sector. Evaluation findings can contribute to developing a future research agenda, leading to evidence-based policy recommendations, and drawing on the findings of a whole body of knowledge *including* evaluations. This positive contribution could be enhanced by strengthening the evaluation skills of social science researchers through exposure to evaluation theories and tools. At the same time, the immensely valuable set of knowledge and experience that has been gained by a number of feminist development practitioners will be unable to reach wider communities of evaluators if their evaluative

⁴ For example, Abhijit Sen, in his keynote address on “The Role of Evaluation in Policy and Programming” at the Evaluation Conclave, Delhi, October 2010, made a strong case for strengthening evaluation as a discipline (Evaluation Conclave 2010).

writings are not framed and structured in the language that is recognized by the conventional evaluators.

Program Objectives

With this understanding, ISST in New Delhi designed a four-year program to address the increasing demand for transparency and accountability from program managers, including the government; the increasing interest in evaluation, but equally recognition of gaps in capacity; and the presence of a group of feminist social science researchers who have also engaged with evaluation.

ISST coordinated the program, which was entitled “Engendering Policy through Evaluation: Uncovering Exclusion, Challenging Inequities,” from 2011 to 2015, in response to persistent gender inequalities in Indian society and the economy in various spheres.⁵ In principle, gender-equitable outcome is regarded as a cross-cutting objective across all sectors of development in India. However, it is often observed that, at the policy level, there is a reluctance to make gender concerns explicit in program design and implementation. It seems policy makers are not convinced whether a gender- and equity-focused lens in program design and evaluation would significantly improve outcomes or not. On the other hand, there is sufficient evidence from existing research that alternative strategies can indeed help in reducing the gender gap. The motivation for the program then was “to try to change and improve things on the ground” by demonstrating evaluation approaches derived from feminist theories of social inequities.

The overall objective of the program was to strengthen the understanding of gendered implications of policies and programs, and to enable the formulation of gender-sensitive approaches. In particular, focus was placed on evaluating selected key issues related to education, health, governance, and livelihood. Specific objectives included building evaluation capacity using a feminist lens, and expanding research and inquiry into the benefits of doing so. The work would contribute to building the field of feminist evaluation, and building an active network of professionals who are engaged in advancing its theory and practice.

KEY ACTIVITIES AND ACHIEVEMENTS

Over a period of four years, this program has built a strong network of individuals and organizations in India that are interested in and working on feminist evaluation (table 12.1). For the purpose of capacity building, expanding a network of persons interested in evaluation with a gender and equity lens, and engaging policy makers, the following activities were undertaken.

⁵ The program was jointly sponsored by the International Development Research Centre, Canada; and Ford Foundation, New Delhi.

TABLE 12.1 **Structure of the program**

Participants	Activities	Outputs
<ul style="list-style-type: none"> ■ Feminist evaluators ■ Gender researchers ■ Members from implementing organizations ■ M&E personnel of different organizations ■ Members from donor agencies/other development agencies 	<ul style="list-style-type: none"> ■ Interactive training workshops ■ Moderation of an online community of practice ■ Participation in evaluation conferences ■ Reflective writings on evaluation experiences with feminist lens ■ Development of resources ■ Dissemination seminars with NITI Aayog 	<ul style="list-style-type: none"> ■ Resource materials on gender transformative evaluations (print and online) ■ Edited collection of meta-evaluations ■ Edited collection of feminist evaluations ■ Short training videos ■ Website and online community of practice

- The network-building activity was sustained through a feminist evaluation Google group; an interactive online community of practice on gender and evaluation; a Facebook page; and a Twitter account.
- The work of capacity building and building a base of knowledge on feminist evaluation was sustained through workshops, small research grants, and support for participating in international workshops and seminars.
- The knowledge produced through this process was disseminated through publications, and both print and online training videos.
- The program has tried to engage policy makers at different stages.

The program conducted six training workshops, which functioned both as capacity-building spaces as well as opportunities to share research and knowledge on feminist evaluation. The program also offered scholarships to attend international conferences.

The program offered 11 small research grants for reflective research on evaluations. Some of the sponsored studies were “A Study of Gender and Water, Sanitation and Hygiene (WASH) Evaluations in India” and “The Culture of Evaluations: Women Empowerment Programs under the Corporate Social Responsibility (CSR) Initiatives.” There were also two studies on meta-evaluations: one in the field of education, the other on health programs of the government of India.

Four books were published as outputs of this program: a toolkit on gender-sensitive participatory evaluation methods; an edited volume, *Engendering Meta-evaluations: Towards Women’s Empowerment*; an edited resource pack, *Resource Pack on Gender-Transformative Evaluations*; and an edited collection of feminist evaluations, *Voices and Values: The Politics of Feminist Evaluation*.

Since the inception of the program, knowledge sharing has been a key component. The base for this was provided by a feminist evaluation website,⁶ various social media pages, and a Google group for those interested in feminist evaluation that was created in the first year of the program, as part of the knowledge sharing strategy. In addition, an online community of practice was set up in early 2013 to build a network of gender-responsive evaluation practitioners and researchers, including the core project participants, but also going beyond this group. The purpose was to provide an interactive, iterative platform that would build a knowledge base on gender and evaluation for the community of practitioners. This online community has more than 3,000 members from all over the world, and has generated new interest in gender and evaluation. It also provided an impetus for several organizations to seek collaborations with our feminist evaluation network. Since 2015, the online platform is also the online knowledge-sharing hub for EvalGender+.

There has been a widening of the networks beyond the initial base of project participants. While the project participants continue to form the core of our network, a broader network has been engendered by the online community, which has provided a much broader base for sustained conversations on gender and evaluation. Training videos on *What Are Gender-Transformative Evaluations?*, *Use of Evaluations in a Gender and Equity Context*, and *Principles, Values and Ethics of Gender-Transformative Evaluations* have been uploaded, which members have used, and through which they have also shared their inputs.

A key component of the program was also to engage policy makers, and to communicate the value of a feminist perspective in evaluations. Attempts to engage the policy makers have been underway since the beginning of the program. However, these efforts did not materialize until year 4, beginning with the International Year of Evaluation (EvalYear) events in India, for which ISST, the National Institution for Transforming India (NITI Aayog),⁷ and the National Institute of Labour Economics Research and Development (NILERD) were co-organizers, along with other organizations.

Armed with some of the outputs of the program that directly address evaluation policies (for example meta-evaluations of government programs and state accountability mechanisms), the process of directly engaging policy makers proved to be more productive. Thus, in the fourth year, the program organized two policy workshops, in association with NITI Aayog and NILERD. And ISST's efforts to engage policymakers to use a feminist perspective in evaluations culminated in the launch of the Evaluation Community of India (ECOI), hosted by ISST.

The capacity to conduct evaluation and recognize the value of evaluations has been strengthened with each workshop, reflection session, and conversation on ISST's online platform. This group has contributed to building

⁶www.feministevaluation.org.

⁷NITI Aayog is a policy think tank established by the Indian government in 2015 to replace the erstwhile government institution called the Planning Commission.

the field in diverse ways, including the publication of papers and making presentations at national and international workshops and conferences, and at capacity-building workshops.

The group has been actively engaging with ongoing policy debates on reproductive and sexual health, particularly through the work on child and early marriages, and adolescent girls; on education (for instance, Sarva Shiksha Abhiyaan);⁸ on the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), through a meta-evaluation study; and on livelihoods, through evaluating the governance of the irrigation sector through a gender lens.

Some of the participant organizations have incorporated their learning from the workshops into their own practice on monitoring and evaluation (M&E). For instance, Jabala, a community-based nongovernmental organization (NGO) that operates from Kolkata, conducted an internal evaluation on their economic rehabilitation program for survivors based on their learning from the workshops. The Child in Need Institute (CINI), an NGO based in eastern India, collaborated with ISST on using community-led participatory M&E tools in their project based in West Bengal. The Centre for Catalyzing Change (formerly known as CEDPA) is keen to develop a self-assessment tool for an end-line survey in one of their projects. And the Health Institute for Mother and Child (MAMTA) has started using participatory evaluation methods in their organization.

ANALYSIS

Given the wide variety of stakeholders that were targeted by the program, the amount and type of involvement, and the takeaways, also varied.

Practitioners were drawn from organizations that were conducting development programs, and were mid-to-senior-level professionals from different disciplinary backgrounds, including social science, management, legal, and others. For this group, evaluation knowledge was rudimentary, and their experience was a “third-party activity” to which they were subjected by donors. Against this background, participation in the project was found to be valuable in enabling ownership of evaluation within the organization: in assisting them to think evaluatively of their work; in unpacking gender dynamics that had been opaque in their program design and implementation; and in showing the way for creating an “evaluative culture” within their organizations.

The program has contributed to many key building blocks toward improving capacity building for organizations, and professionals engaged in development work at the grassroots level. Such organizations typically rely on third-party evaluations, and do not engage in viewing their own work using a feminist lens. This was the case even for organizations that were working

⁸ Sarva Shiksha Abhiyan (SSA) is Government of India’s flagship program for achievement of Universalization of Elementary Education in a time-bound manner, as mandated by the 86th Amendment to the Constitution of India, making free and compulsory education to children in the 6–14 age group a fundamental right.

on equity and women's rights issues. Such an engagement was found to be valuable in identifying dimensions for engagement, or for ongoing analysis of their work. Thus, feminist evaluation concepts of power and structures were becoming integrated into ongoing and new projects, and these opened up new ways of looking at empowerment. This capacity is crucial in order for an organization to actively engage in ongoing self-evaluation that can feed into the design of more formal periodic evaluations. The capacity-building workshops enabled development practitioners to better understand and measure change processes, and reflect on their own experiences.

Capacity building for conducting formal self-evaluations by these organizations was found to be more limiting, especially as there tended to be rotation among the staff who attended, and the workshop sessions were not geared to be evaluation training per se. With the evaluative thinking on feminist issues that was provided by the program, individual organizations embarked on their own evaluations, with mentoring support from experienced evaluators, and also with the support of ISST.

More senior practitioners, such as those who were involved with joint review missions for the government's flagship programs, found the feminist evaluation concepts very helpful in assisting them to formulate questions that helped to push the analysis beyond the data that was available for these programs. Integrating these ideas with their government counterparts was also well received.

Researchers and evaluators working on feminist evaluation found the dynamic interaction on research issues very helpful in furthering their work. Such opportunities for reflection, interaction, and feedback are valuable for conducting good research.

The researchers who participated in the program were at both the junior and senior levels. The junior-level researchers valued the ability to conduct work and the opportunity to attend and present results at international conferences provided by the research grants, and the reviews and discussions that enabled them to complete their publications. The senior researchers, who had taken on higher-profile meta-evaluations of national flagship programs, were supported for conducting and publishing the work in an edited volume, which subsequently has been released at a policy seminar cohosted by NITI Aayog. Being able to influence policy makers was enhanced by the involvement of other experienced evaluators who were also involved in this program. The series of workshops provided a space for reflection and introduction to innovative ideas that spurred research productivity. And the e-network provided an ongoing space for discussion of issues, concepts, and methods, and for researchers to get feedback on their work.

There was relatively little engagement with decision makers and policy makers from institutional settings where programs were being designed and/or implemented; where evaluation training was being carried out; or with government officials. The concept underlying this work was that results emanating from the capacity building and research output would subsequently be able to influence decision makers.

The program outputs have the potential to influence many of these groups of decision makers. For example, the training modules and resource

pack could contribute to the integration of these methods in standard evaluation trainings; and research outputs such as the compiled meta-evaluations of national flagship programs could influence a closer look at their implementation and modifications. The dissemination and outreach workshops and seminars conducted in the final year of the program highlighted the fact that such influence was indeed being generated.

Organizational leaders wanted more hands-on training in order to be able to conduct their evaluations. The program has introduced fresh perspectives to work being done by the NGOs that participated, and evaluative thinking was being integrated into their programs. However, they lacked the capacity and expertise to conduct their own evaluations, and would have liked to have an ongoing mentoring relationship established to facilitate that.

Evaluation practitioners, especially those who were active in mainstream program evaluations, had a mixed response. Some responded that the reflection and research conducted with their participation had enabled them to better interject feminist issues into program designing and implementation, though they felt that the terminology of gender-transformative evaluation was more acceptable within their own constituencies. Others felt constrained with their ability to apply feminist evaluation concepts in their practice.

Some of the reasons they cited for this included:

- Commissioners of evaluation, and most donors, pay only lip service to gender empowerment/transformational development;
- The need to demonstrate links to economic productivity with equity/gender empowerment programming; and
- Logistical issues in conducting fieldwork by female evaluators for getting women's perspectives.

The volume of high-quality research output and publications with international visibility has contributed to energizing this field of inquiry: a significant majority of participants said that this was a valuable contribution, and that a desire for ongoing engagement with the community of practice has been established.

PROGRAM OUTCOMES AS A WHOLE

Short Term

The regular workshops conducted over the course of the four years of program implementation created a vibrant space for interaction, reflection, and the sharing of innovative methods and approaches that engaged participants in a productive manner. Those who attended several of these events gained insights that they introduced to their organizations and into their work. The shared learning space and commonality of equity and feminist program and research interests of the participants was conducive to creating an effective community of practice.

The external visibility produced both within the participants' own organizations as well as in international forums, is likely to enable feedback loops that will contribute to benefits in the medium and longer term.

Medium Term

Publications of research on feminist evaluation include one edited collection of meta-evaluations of government programs (Sudarshan, Murthy, and Chigateri 2015), one toolkit on gender-sensitive participatory evaluation methods (Murthy 2015), one resource pack on gender-transformative evaluations (Chigateri and Saha 2016), and one edited volume on reflective writings on the practice of feminist evaluation in India (Sudarshan and Nandi 2018). These publications will provide an impetus for additional research and integration of a feminist lens in evaluations, program reviews, and their design.

The e-network, established and expanded internationally to more than a thousand members during the program period, is expected to be a major instrument for providing an open forum and long-term benefits for the continued development in the field of feminist evaluation that was spurred by this program.

Long Term

The yardstick that can be used to assess the longer-term outcomes of this program is the extent to which there is institutionalization of the key short and medium-term outcomes. The main evidence of this is from interviews with NGO practitioners who are integrating evaluative thinking into their programs, and who have acquired the capacity to apply a feminist lens to their programs. Those working in more mainstream programs did not face such a positive institutional environment.

The program was able to develop some integration of feminist or gender-transformative lenses in the evaluation training being offered at NILERD, which partnered in some of the workshops and outreach events. With the dissemination of publications that resulted from the program, there is potential for expanding such capacity building.

The impressive volume of high-quality research, and the publications that have been produced, will also contribute to longer-term outcomes.

SUSTAINABILITY OF THE PROGRAM

Human resource capacities built at the individual level by the program will be sustainable, given the depth and diversity of the discourse that has been provided. It is anticipated that these individuals will carry this capacity into their ongoing work, since they have been enabled to establish this type of discourse into any of the programs, evaluations, or research contexts in which they may find themselves.

Since the program was designed to influence capacity at the individual level only, at the organizational level influence is most likely to occur as the indirect result of participants who are able to integrate their learning on the

value and processes of introducing a feminist lens into the work of their organizations. As discussed earlier, this type of institutionalization was more apparent in the women's and human rights-based organizations, where the shift was related to introducing access to evaluative thinking, and a feminist lens enabled them to identify dimensions of their work that were not evident earlier.

However, the senior participants from mainstream development organizations felt confident that they would be able to make some inroads into the thinking at the institutions they worked with. At the same time, they also identified the constraints they faced. One of these is the challenge of using the terminology. Often the application of the term "feminist lens" has a tendency to address gender issues in a superficial manner, with little or no demand for systematic, gender-transformative programming and policy formulations, or their evaluation. And a number of the junior-level professionals did not feel competent enough to influence the work in their respective organizations.

The program has managed to have a small amount of direct engagement with national and state government officials concerning the value of using a feminist or gender lens in program evaluation, and in addressing program design for equity outcomes. This engagement only happened in the final year of the program, and was primarily the result of a few workshops and seminars that were cohosted with NITI Aayog. The previous chief executive officer of NITI Aayog and a number of senior-level officials participated, and engaged actively in those events. The program results presented were received very positively, and a favorable policy environment was evident. However, for this to be sustained and integrated into mainstream program evaluations and design will require concerted, ongoing work.

CONCLUSIONS AND RECOMMENDATIONS

Participants from rights-based organizations found value in learning from the program and the workshops. Many of them who are new to evaluation began to integrate evaluative thinking in their work.

The program provided a new way of looking at how to make evaluation gender responsive. But it was challenging at the organizational level, since new learning applications had to go through an internal process of explaining and convincing the leadership. Introducing a gender-transformative perspective in designing and implementing a program even at the organizational level would be slow and challenging.

Researchers opined that the program enabled them to connect with practitioners. As for using the learning in their work, a couple of them noted that the term "feminist lens" was not always acceptable to their colleagues, so they used the concepts but not the terminology. The online community was very helpful, and much was learned from the information sharing.⁹ Overall, there has been a positive contribution to building the field for feminist/

⁹The online community of practice can be accessed at <http://genderevalning.com>.

gender-transformative evaluation and opinions on transforming policies and programs using gendered evaluation.

However, there are serious challenges in conceptualizing feminist evaluation. One major challenge has been confusion around the question of whether equity-focused and gender-responsive evaluation is different than feminist evaluation. The question was raised time and again, both inside the group and outside of it, whether feminist evaluation is a separate methodology or simply a frame of mind to use in looking at issues of inequality and gender-based power dynamics. There is also confusion as to whether feminist evaluation is an approach that could be used in any evaluation practice. This was discussed in some of the group workshops: whether the feminist approach should be integrated into the project design, or not. The question of whether to use a feminist lens in project monitoring was another area of concern.

During the midterm evaluation workshop, some of the participants raised the issue of wider acceptance of evaluation reports conducted with a feminist lens. One participant raised her concern that policy makers do not want complexity: they like to see a few clear results, stated in terms of numbers. This may go against the grain of feminist evaluation. The opinion was expressed in this workshop that the feminist evaluation approach cannot produce figures and numbers: rather, it captures social changes or program/project outcomes qualitatively. Perhaps this is the reason that policy makers and governments rely mostly on a set of institutes who practice and adopt more quantitative methods.

This brings up the second level of concerns for the group of feminist evaluators. The question was asked, how to bring feminist values into the evaluation findings, and how these values can be flagged so that policy makers will start noticing them. Other questions discussed were how to create an enabling environment for equity-focused and gender-responsive evaluation, and how to share lessons learned effectively in lessening gender inequalities. On this last question the group agreed that practicing feminist evaluation is part of a larger structure of feminist politics that ultimately tries to bring equality. This concern enabled the group to name feminist evaluation practice as gender-transformative evaluation in an Indian context.

During these interactive sessions, it was understood that there is a need for further dialogue and discussion among the members of the group and among both young and senior practitioners and evaluators. A number of participants in the program were open to sharing and exchanging. The recognition of the importance of converging various approaches and methods even increased greatly with the progress of the program. The group agreed on a crucial issue that feminist evaluation knowledge needs to go beyond feminist constituency and reach a wider audience. The group also discussed how to deepen the alignment between feminist researchers, mainstream evaluators, and formal and informal networks of evaluators, donors, university and research institutes, governments, and NGOs in order to build the field of feminist evaluation.

During the midterm external review workshop in 2014, the group agreed to strengthen the capacities of a larger group of development

practitioners and M&E personnel through developing an easy-to-read toolkit and the mentoring of young practitioners by senior feminist evaluators. They also suggested that the community of practice can facilitate a space where people will be able to interact, debate, discuss, and share knowledge.

A feminist evaluation curriculum for India and South Asia that is mainstreamed through the university system would help in developing capacity in feminist evaluation. To ensure that feminist evaluation is adopted at the organizational level, there is a need to orient staff, particularly the leadership. The group also felt a need to engage more with government officials, and with a focused and targeted approach.

A series of suggestions was rendered by the evaluators at the end of the four-year program. These suggestions came out through conversations with members of the group. One important suggestion was to continue organizing policy workshops, publishing policy briefs, and publishing in reputable journals and edited volumes in order to document the process of change that feminist evaluation has created. The second important suggestion was to find ways to increase visibility of the group of feminist evaluators in India. The third suggestion was to make efforts to integrate equity-focused, gender-responsive evaluation with mainstream evaluation methods such as utilization-focused evaluation, or impact evaluation.

In the longer run, organizational capacity building for integrating a feminist lens would be aimed at the foundation that has been built by this program. Attention is needed to support continued research on documenting the approaches and the value gained with feminist evaluation, and its integration into mainstream development research agendas.

Continuing the feminist evaluation e-network and community of practice is a priority. The recent development of EvalGender+ stepping in to support this is a big plus, and will help to maintain the momentum that has been generated. This platform has been valuable for encouraging exchange on theory and practice, creating a space in which to question and have a dialogue on issues, challenges, different methods, and their applications.

Outreach and dissemination of key messages and lessons learned from the knowledge products that have been produced for different audiences will help to expand the audience and understanding gained beyond those who are active participants in this work.

As a follow-up of the program, the core group of feminist evaluators might consider introducing courses in feminist/gender-transformative evaluation at evaluation training centers in India. Gender is becoming a standard cross-cutting objective in development programs, but not enough attention is being given to how the objectives can be attained in reality. At the same time, building of capacity at the organizational level for integrating a feminist lens in the design and monitoring of programs for producing gender-equitable results will be another key step. This can be linked to working with donors to promote building a culture of evaluation with a feminist lens.

The core group of gender evaluation network participants, along with ISST, are well positioned to expand collaboration with NITI Aayog and the state governments to review their evaluation guidelines and methodologies, and their M&E review processes. In some states, leaders of this feminist

evaluation network are already assisting with reform of the M&E processes, and are documenting their work so that it can be used to assist in cross-learning mobilizing support.

This would also be an interesting way to get the insights and involvement of the leaders in feminist evaluation, who otherwise do not have time to contribute to the online platforms or join the discussion groups.

Research funding for feminist evaluation is an ongoing constraint, and the online community of practice can take up this issue and find ways to establish a research fund for supporting ongoing research.

ACKNOWLEDGMENTS

The authors would like to acknowledge that the chapter is based on various project documents. A number of team members were involved in drafting those documents. The authors would like to mention particularly Katherine Hay and Ratna M. Sudarshan for conceptualizing the project along with a core group of feminist researchers and evaluators, and for developing the initial project documents and research papers. Two evaluation reports that were used as a source for this paper were drafted by Kalyani Menon Sen and Shubh Kumar-Range. The authors are grateful to all members of the feminist evaluation group for their contributions in the project and to the discussions. Finally the authors gratefully acknowledge the sponsoring organizations: the International Development Research Centre, Canada, and the Ford Foundation, New Delhi.

REFERENCES

- Bakshi, M., and M. Bhattacharya. 2010. "Miles to Go: Effectiveness of RTI for Women." *Global Media Journal—India Edition* December.
- Chigateri, Shraddha, Tanisha Jugran, Rituu B. Nanda, and Rajib Nandi. 2015. "Learning from STEP Evaluations." In *Engendering Meta-Evaluations: Towards Women's Empowerment*, Ratna M. Sudarshan, Ranjani K. Murthy, and Shraddha Chigateri, eds. New Delhi: Institute of Social Studies Trust.
- Chigateri, Shraddha, and Shiny Saha. 2016. *A Resource Pack on Gender-Transformative Evaluations*. New Delhi: Institute of Social Studies Trust.
- Evaluation Conclave. 2010. "The Evaluation Conclave 2010: Making Evaluation Matter. Communication Outreach Report." New Delhi.
- GOI (Government of India). 2016. "Women and Men in India 2016." Ministry of Statistics and Programme Implementation, New Delhi.
- Hay, Katherine, et al. 2012. "Why a Special Issue on Evaluating Gender and Equity?" *Indian Journal of Gender Studies* 19 (2): 179–86.
- Murthy, Ranjani K. 2015. *Toolkit on Gender-Sensitive Participatory Evaluation Methods*. New Delhi: Institute of Social Studies Trust.
- Planning Commission. 2011a. "Faster, Sustainable and More Inclusive Growth: An Approach to the Twelfth Five Year Plan." Government of India, New Delhi.
- . 2011b. "Issues for Approach to the 12th Five Year Plan." Presentation, UNDAF Orientation Workshop, New Delhi, May 12, 2011.
- . 2011c. *Mid-Term Appraisal: Eleventh Five Year Plan 2007–2012*. New Delhi: Government of India and Oxford University Press.

- Shiva Kumar, A.K. 2010. "A Comment on 'Evaluation Field Building in South Asia: Reflections, Anecdotes and Questions.'" *American Journal of Evaluation* 31 (2): 238–40.
- Sudarshan, Ratna M., Ranjani K. Murthy, and Shraddha Chigateri, eds. 2015. *Engendering Meta-Evaluations: Towards Women's Empowerment*. New Delhi: Institute of Social Studies Trust.
- Sudarshan, Ratna M., and Rajib Nandi, eds. 2018. *Voices and Values: The Politics of Feminist Evaluation*. New Delhi: Zubaan.
- UNICEF (United Nations Children's Fund). 2014. *Ending Child Marriage: Progress and Prospects*. New York: UNICEF.
- WNTA (Wada Na Todo Abhiyan). 2011. *Approaching Equity: Civil Society Inputs for the Approach Paper—12th Five Year Plan*. New Delhi: WNTA and United Nations Development Programme, India.
- Young Lives. 2016. "Tackling Child Marriage and Early Childbearing in India: Lessons from Young Lives." Policy Brief 29. Oxford Department of International Development, Oxford, UK.

Chapter 13

Evaluation Cooperation in West Africa

Abdoulaye Gounou

Abstract. *Evaluation is an increasing concern for the francophone countries of West Africa: however, for the most part impact assessment does not yet have the interest it should in these countries. To reverse this trend, under Benin's leadership, a capacity-building program has been initiated to promote impact evaluation as a tool for public policy analysis, support impact evaluation studies in these countries, and gain consideration for the results of impact evaluation within the framework of public management. This innovative project involves South-South cooperation between countries that share roughly similar economic, political, and social contexts, as well as a legal and economic framework that is converging. The growing interest of West African states in evaluation has led several countries to develop their own evaluation systems. In 2012, eight countries with similar interests met for a workshop on monitoring and evaluation. As a result of the workshop, three countries—Benin, South Africa, and Uganda—are continuing to cooperate by developing the Twende Mbele ("Let's All Move Forward") program to strengthen performance and evaluation monitoring. This program aims to strengthen national evaluation systems, and to gradually extend its interventions to other African countries through the development of appropriate tools for monitoring and evaluation, the capitalization of knowledge resulting from evaluations, and the sharing of national experiences in evaluation.*

Francophone West Africa is identified as an area where there is a need for capacity building in evaluation. The practice of evaluation is itself very uneven and, depending on the country, the institutionalization of evaluation is still rudimentary. This observation was made during the francophone dialogue on evaluation capacity development held in Cotonou on July 7, 2015. More specifically, in the countries of the West African Economic and Monetary Union (WAEMU), the production and use of evidence through impact evaluations to inform public decision making is still very weak.¹

Several initiatives at the national and regional levels have been undertaken to provide answers to this unpleasant picture as it has been observed in francophone West Africa. These initiatives include national evaluation days as well as regional programs such as Twende Mbele (“Let’s All Move Forward”) and the Capacity Building and Impact Assessment Program in West Africa (WACIE).

This chapter will show how such initiatives can help to effectively strengthen the institutionalization of evaluation, and improve evaluation practice in the francophone countries of West Africa. It will also show that the development of evaluation in these countries has also brought about the design and implementation of regional programs for the sharing of best practices, and support for sustainable capacity building.

The chapter is structured in four parts, plus a conclusion. The first part describes Benin’s experience, in particular progress made in the field of evaluation and how cooperation has influenced this evolution. The second and third parts, respectively, present the WACIE and Twende Mbele programs. The fourth part demonstrates how effective implementation of these two programs and their interrelationship offers clear opportunities for the development of evaluation at the regional level.

BENIN’S EXPERIENCE IN REGIONAL COOPERATION

In 2007, Benin initiated the process of developing evaluative work by assigning this task to a ministry. Within the framework of the operational management of this function, the Office of Public Action Evaluation; the Bureau of Public Policy Evaluation; the Directorate General of Evaluation; and finally (and currently), the Bureau for the Evaluation of Public Policies and Analysis of Governmental Action were established.

The strategy adopted by this institution from its establishment in 2007 until the present basically involves benchmarking and integrating Benin into international evaluation networks. By participating in international evaluation conferences, Benin realized that South Africa and Uganda are in the lead in establishing national monitoring and evaluation (M&E) systems. In 2012, the government of Benin endowed the country with a national evaluation policy (NEP 2012–2021), and set up an institutional framework for public policy evaluation, which includes the organizations in charge of M&E and development planning.

¹For more information, see Mendiratta (2011).

These advances at the national level have resulted in closer cooperation between Benin, South Africa, and Uganda, and as they appreciated the country's efforts in this area and its lead over other francophone West African counterparts.

Determination from its national stakeholders, supported by strong political will, has enabled Benin to join its peers from South Africa and Uganda in the evaluation cooperation program called Twende Mbele.

As members of the International Initiative for Impact Evaluation (3ie), South Africa and Uganda encouraged Benin to join this initiative for help in developing the practice of impact evaluation. In 2014, Benin signed a memorandum with 3ie that enables it to receive annual grants for the commissioning of impact assessments, and for capacity-building activities in impact evaluation.

When participating in 3ie's Evidence Weeks, Benin observed the absence of almost all the francophone countries of the WAEMU area. The initiative for a regional development program of impact assessment in the WAEMU area was designed, and eventually evolved into WACIE.

3ie's traditional field of intervention is impact evaluation, but the institution has broadened its scope to include the area of monitoring. Impact evaluation remains the organization's main focus, but a window is now open for providing support for the strengthening of national M&E systems, and the promotion of researchers from Global South countries, in an attempt to help reduce the gap between North and South in the field of evaluation.

Especially in West Africa, national evaluation systems are weak, the culture of evaluation is still in an early stage, and the practice of impact evaluation is almost nonexistent. This weak culture is caused by a series of difficulties that the promotion and development of evaluation faces, in particular low demand for evaluation in the region; weakness of national statistical systems; and the minimal degree of ownership of evaluation issues by states that have left the demand for and funding of evaluations to donors.

WACIE offers opportunities in the West Africa region for collaboration with the Twende Mbele program, which is a continental initiative. Benin's leadership in promoting evaluation has been demonstrated through its official support for WACIE, and its support for French-speaking countries in West Africa to look to Twende Mbele to help them make progress on governance and accountability issues. Benin promotes the Twende Mbele program in West Africa. And WACIE, which is a regional initiative, has become an opportunity for Twende Mbele to reach many more audiences.

THE CAPACITY BUILDING AND IMPACT EVALUATION PROGRAM IN WEST AFRICA (WACIE)

It is now recognized that rigorous impact evaluations can help quantify the socioeconomic impacts of projects and programs in several areas, and can provide a solid basis for consistent policy-making decisions. Although there is a growing global interest in impact evaluations, a report on trends in Sub-Saharan Africa revealed that there is little to no buy-in from countries and nationals of the continent (Mendiratta. 2011). Despite a considerable

increase in terms of the number of evaluations since 2004 (77 percent from 2004 onward), there are still significant disparities. For instance, impact evaluations in Africa are conducted much more frequently in anglophone countries (mainly Kenya, followed by Uganda) compared to francophone countries (18 percent).² In addition, only 11 percent of the studies³ are conducted with African experts involved in the drafting of the research paper, although African partners (e.g., local NGOs, ministries, etc.) are involved in various stages of the implementation of the programs.

Over the past two years, discussions with key regional stakeholders, including the African Development Bank, the West African Development Bank, and the governments of Benin, Guinea Bissau, Niger, and Senegal have identified a real need for capacity building in order to produce high-quality research results, and to promote their use in the development and implementation of policies. While most governments in the West African region have monitoring systems in place, three key conditions limit rigorous impact evaluations: a lack of the human resources and institutional capacity that could help to conduct and use impact evaluations; a limited monitoring system for helping in the implementation of evaluations; and a low level of competence to conduct the required studies. WACIE was created to tackle these issues.

WACIE is a regionally based national initiative to improve welfare and development outcomes through decision making that is informed by research findings in the West African region. The initial phase of the program will begin in year 2017 and will last three years, followed by a consolidation phase, which may be carried out by a regional organization within WAEMU. The program is promoted by the government of Benin, with partial technical and financial support from 3ie.

Key Elements of the Program

As specified in the program documents, WACIE's overall objective is to improve welfare outcomes through informed decision making, based on evidence from research findings in the region. To this end, the program is characterized by several key elements:

- Initiating and managing consultations at the national level to identify key areas where evidence is needed for policy making and programmatic decision making
- Establishing and managing a funding mechanism to sponsor public policy impact evaluations
- Developing a set of standards, guidelines, and databases to guide and communicate evaluation results
- Designing, implementing, and disseminating assessments

² Angola, Cape Verde, Eritrea, Mauritius, Mozambique, and Rwanda are classified as neither English- nor French-speaking countries in the report. They account for about 12 percent of impact assessments.

³ The statistics are based on 257 evaluations from the years 1982–2010.

- Providing training and capacity building to support the planning and implementation of impact evaluations, and the use of evidence in the decision-making process in the region

The proposed initiative covers a number of areas.

- **Capacity strengthening to produce and use research findings.**

- The capacity to produce scientific evidence will be enhanced through the participation of local researchers in the funded studies.⁴ Training in impact evaluation will also be offered within the member countries, in collaboration with the Regional Centers for Learning on Evaluation and Results (CLEAR).
- The creation of partnerships, codevelopment, and coeducation of training programs with regional academic institutions such as the African School of Economics.⁵
- The capacity to use research results will also be developed through one or more workshops targeted at decision makers that will be organized within the country.
- The provision of grants to provide staff for trainings and conferences.
- Support through an assurance of assistance for quality impact evaluation of local governments and partners in the region.
- Development of a network to connect personnel, researchers, and decision makers in the region. Furthermore, study teams will be made available to conduct workshops with their respective implementing agencies, thereby enhancing impact evaluation know-how and the use of research findings in institutions for the conception and implementation of projects and programs.

- **Creation of a database of research results.** This will involve support for at least four impact assessment studies. These studies will focus on areas and research issues identified by WACIE member governments in West Africa. They will be based on consultation sessions supported by WACIE secretariat staff, and supervised by 3ie staff. Staff and study teams will liaise with line ministries to discuss opportunities for impact evaluation of selected programs and their implications, and 3ie will conduct the studies independently. The conception phase of the impact evaluation will be subject to formative evaluation of the implementation process.

- **Policy interpretation of research findings** 3ie has a well-developed system for supporting and monitoring the use of study results. Through this system, the project will support the production of the

⁴This is a condition that must be fulfilled in order to receive funding from 3ie.

⁵The African School of Economics is a private pan-African research university based in Abomey-Calavi, Benin.

operational synthesis of results, and one-page research recommendations on all studies, along with other guidelines derived from research findings that are relevant to help develop major policies.

- **Learning from international experience.** The program will include a grant component to enable participation in international conferences and workshops on impact evaluation, including preconference training. A total of 16 grants will be offered each year.
- **Program management.** The program will be managed by 3ie international staff based in New Delhi, supported by a secretariat located in Cotonou.

Logic of the Intervention

By the end of its implementation, the program is expected to contribute to improving the living conditions of the people of the WAEMU region through more effective public policies. The expected effects are:

- Improving the effectiveness and relevance of national and sub-regional projects and programs;
- Making use of systematic and sustainable evidence in the formulation and implementation of public policies; and
- Ensuring effective program management.

To observe these effects, the following causal pathways should be achieved:

- Evidence is produced and used to influence and support the formulation and implementation of public policies in each country, as well as subregional programs in the WAEMU area
- The WAEMU countries' national evaluation systems are strengthened and able to produce evidence for informed decision making

The intervention of WACIE is therefore designed to produce evidence that can highlight public policy choices in francophone West Africa, with a view toward contributing to the ultimate improvement of the living conditions of the people of the WAEMU area. We are, however, assuming that evidence may not necessarily lead to policy changes for various reasons: for example, because other policies have priority; or because doubts are raised about the evidence; or because trade-offs have been made to achieve political compromises.

In this respect, some of the risks are related to the lack of will of the decision makers or the political context in certain countries.

To lessen these risks, it is necessary to ensure the commitment made by all WACIE member states when they join 3ie. This commitment stipulates, among other things, that the member countries of 3ie will commit to using the results of the impact evaluations financed by 3ie in their countries, in order to improve the quality of public policies. Similarly, the ministers responsible for public policy evaluation in each WAEMU country will be involved in

defining the program's orientations, and will avert risks to the judicious use of the results and recommendations of evaluations.

To support the planning and implementation of impact evaluations and the use of evidence generated by such evaluations, WACIE plans to strengthen the technical capacities of the various stakeholders. Scholarships for training in impact assessment, or participation in technical workshops on impact evaluation in the subregion, will be awarded for this purpose. This ambition could be undermined by the lack of a training center or high-level technical workshop on impact evaluation in the subregion. For this reason, the program will support and encourage institutions and research centers that have the capacities to offer quality trainings or technical workshops on impact evaluation, to increase their offerings.

Milestones and Stakeholders

Six of the eight WAEMU countries (Benin, Burkina Faso, Côte d'Ivoire, Guinea, Niger, and Senegal), and the West African Development Bank, are now members of 3ie. 3ie has also already supported nine impact assessments (in Benin, Burkina Faso, Mali, Niger, and Senegal). These impact assessments were preceded by capacity-building workshops and consultations. 3ie has also supported four capacity-building workshops in the region (in Benin, Burkina Faso, and Togo).

WACIE will bring together as stakeholders the member states of WAEMU; WAEMU regional institutions (i.e., the WAEMU donor agencies); 3ie; and other nonregional donors (i.e., donor agencies outside of WAEMU), as well as other stakeholders.

- **WAEMU member states.** Each country participating in WACIE will be identifying and choosing a focal point for working with the WACIE secretariat. Focal points are public institutions with a legal mandate to work on impact evaluation issues with the support of their respective government authorities. The role of the focal points is to facilitate the implementation of WACIE at the country level: they will play a key role in the consultation process of identifying relevant projects in relation to the socioeconomic context of each country. In addition, they will ensure the effective involvement of local and national institutions in the initiative.
- **WAEMU regional institutions (WAEMU donor agencies).** These are the West African Development Bank and the WAEMU Commission. They represent WAEMU internal donors who will be supporting WACIE and helping to define strategic policy by participating in WACIE Advisory Committee.
- **3ie.** 3ie will act on the one hand as a catalyst throughout the initiative, by motivating regional involvement and participation in the program; and on the other hand, coordinating and supporting some WACIE activities. WACIE will be led by 3ie's main office in New Delhi. The design of WACIE assumes that there are resources for impact evaluation in the West African region, and that these resources need to be channeled through training and technical assistance.

- **Other nonregional donors (non-WAEMU donor agencies).**

These are other technical and financial partners from the WAEMU member states that will cofinance the program: for example, the Agence Française de Développement (AFD), the African Development Bank, and others. They will help mobilize resources and define WACIE's strategic policy by participating in the its Advisory Committee. WACIE support modalities are presented later in this chapter.

- **Other stakeholders.** This group includes the project-executing agencies that will be selected under WACIE. It also includes research teams, and all institutions not included in the previous groups that will benefit from or contribute to the initiative.

STRENGTHENING PERFORMANCE AND EVALUATION MONITORING PROGRAMS: TWENDE MBELE

In March 2012, The South African Department of Performance Planning and Monitoring and CLEAR in anglophone Africa organized a regional workshop with seven African countries in Johannesburg in March 2012. At this workshop, Benin, South Africa, and Uganda recognized the similarity of their approaches to developing their national evaluation systems. Since then, they have collaborated by sharing guidelines and methodological tools, and attending training and events together. This has allowed for some cross-learning and the building of a slight, but effective, collaboration between the three nations. And in both anglophone and francophone Africa, CLEAR has worked with these three countries during the initiative.

Since 2013, the United Kingdom's Department for International Development (DFID) has been supporting the South African Department of Performance in their Strengthening Performance M&E project, which is focusing on the use of M&E in front-line service delivery and on citizen monitoring to inform the government on results delivered to the people of South Africa. The result of the project is that effective delivery of front-line services, a citizen-based surveillance system, and a results-based evaluation system support government accountability to the South African populace. A total of £2 million has been made available over 2013–15. The project has been successful: the systems have developed rapidly and have had impacts on the formulation and implementation of policies. This South African experience has been shared as a training in the framework of cooperation between Benin, South Africa, and Uganda.

These three countries have been active promoters of regional and international networks, including CLEAR and 3ie; both are supported by DFID. South Africa and the United Kingdom are also members of the Partnership for Open Government, which promotes accountability and transparency in government action. Other countries have been interested in what these three countries are doing, with study visits from countries such as Botswana, Kenya, Lesotho, Mauritius, Niger, Seychelles, and Togo. They are sharing their experiences and tools widely in the region, both in evaluation and monitoring, especially when they rely on the regional work CLEAR undertakes.

The Twende Mbele program aims to support emerging countries across Africa, enabling the three founding partners to intensify learning and to involve other partners that are determined to move forward with M&E. The program will also be built on global experience, using a range of M&E mechanisms that have been proven to improve government performance and accountability.

DFID is currently funding the African Networks Strengthening for Governance, Accountability, and Transparency (SANGAT) program, which supports three other African networks. SANGAT aims to take advantage of the existing projects, demand-driven processes, and peer learning in thematic areas. There are currently three SANGAT projects, focusing on areas from budget transparency and public finance to a high-level network dealing with major issues of prosecution and organized crime. Twende Mbele is being funded as a fourth component of the SANGAT program.

Key Elements of the Program

As specified in the project documents, the immediate result proposed for Twende Mbele is “the improvement of M&E systems (e.g., practices, policies, tools and procedures) demonstrated in partner countries based on shared experiences.” The theory of change assumes that the M&E systems demonstrated in the project (immediate outcomes) will be expanded and further developed, and will be applied in additional countries over a period of six years (wider result) in order to improve the output and accountability of governments toward their citizens. This would result in improved services, a better use of M&E products by parliaments, and an improved performance culture.

The immediate results—demonstration of improved M&E systems—will be achieved by:

- The creation of an M&E application at the level of senior management, Parliament, and the public;
- The conduct of learning and sharing activities to build on, based on the experience of current partners;
- The development of specific M&E tools in collaboration with partners; and
- The implementation of programs to ensure ownership and cost-effectiveness.

Many countries undertake surveillance activities, but this program seeks to work specifically with those seeking to make a systematic assessment within government as well as monitoring, so that there is a desire for learning and deeper change, and more systemic effects on the output and impact of governments on the living conditions of their people.

Milestone and Stakeholders

The central objective of Twende Mbele is not to constitute a network where countries share their experiences a priori, but to help collaborate in the development and implementation of M&E systems that improve the performance and impact of government actions on citizens.

Established in 2015, the plan predicts that by 2018, the number of partner countries will increase from three to six. There is already high demand for countries to be included in the project, but participating countries must maintain a commitment level corresponding to their capacities. It is therefore important to define an engagement strategy to include other countries. For those countries that do not have a sufficient level of development of their national M&E systems, Twende Mbele has put in place an inclusion mechanism based on differentiations between countries. This option works primarily with countries that are truly committed to using M&E to improve the lives of their citizens.

An initial classification has placed countries in three broad categories:

- **Category A.** Those “knocking on the door” are committed to using M&E in change processes (the indirect indicator being that they are engaged in the evaluation)
- **Category B.** Those who are already initiating actions in the field of M&E, have individual champions, and are willing to become more involved
- **Category C.** Those with some interest but who are not actively pursuing M&E as a key element in improving government performance

The program wishes to actively target all of these countries, and include them to the extent possible in its activities. Their involvement and ability to advance the work will be evaluated using explicit criteria, and enabling them to become collaborative partners.

Potential partners in **Category A** are already very active in M&E, and are keen to improve what they are already doing for better government performance. In these countries, there is already the political will to expose failures and draw lessons from them. These partners could improve their activities by appropriating and contextualizing some of the more advanced tools that have been operationalized in countries outside of Africa. The most effective strategy to employ with these countries is to collaborate on the development of their national M&E systems in order to make them more effective and inclusive. This category includes Côte d'Ivoire, Ghana, and Niger. They can be considered main partners in the program.

The countries in **Category B** have some involvement in M&E, but have not yet committed to extensive systems, or faced rigorous evaluations that may reveal weaknesses. For those with significant interest in M&E, participation in learning and capacity-building activities can be beneficial. However, these countries need to demonstrate their interest in using these opportunities. This category includes Botswana, Ethiopia, Kenya, Lesotho, Rwanda, Senegal, Tanzania, and Zambia.

Category C countries may not be interested in being involved in Twende Mbele, but it may be appropriate to involve them in activities such as newsletters, and informing them of events they might be able to attend. The program should aim to spread M&E focal points across all African countries, and to share news about Twende Mbele as well as general M&E activities

and information that might stimulate further interest. This category includes Angola and the Democratic Republic of Congo.

It is essential that this process be based on self-selection, and that countries recognize where they stand in relation to Twende Mbele. They must also commit to the appropriate roles for each level of participation if they wish to participate.

LEARNING FROM THE BENIN EXPERIENCE

Becoming a member of 3ie has allowed Benin to benefit from financial and technical support for undertaking the impact evaluation of the Free Support for the Scholarship of Girls in Secondary School Level One Program. This program is designed to promote the schooling of young girls across the country and contribute to reducing the education gap with young boys.

The government of Benin is waiting for the evaluation findings in order to improve its sectoral reforms in the field of education. 3ie has also given financial support to many Beninese in the field of capacity building, to attend courses and seminars around the world that are linked to impact evaluation.

Difficulties Encountered

The major difficulty faced up to now in the area of cooperation to promote M&E—both at the country and subregional levels—is related to WACIE, which, while it is a country-level initiative, has program objectives that remain regional.

In fact, the promotion of WACIE requires WAEMU country members to belong to 3ie. Each WAEMU country member has the responsibility to fulfill this essential condition in order for the program to become a reality. Program implementation needs the strong engagement of all of its stakeholders, especially the WAEMU states, and communitarian institutions such as the WAEMU Commission and the West African Development Bank.

Perspectives

Benin's approach to its evaluation system has enabled the country to build its international reputation for the innovative cooperation programs it has developed with its partners. The ultimate expected result of the Twende Mbele program is the effectiveness of national M&E systems at the continental level: WACIE, which is a regional community program, offers promising prospects for the development of Impact evaluation at the regional level.

The sharing of experience and the development of collaborative governance tools encouraged by these programs will undoubtedly improve the quality of transparency in public management at the continental level.

After its pilot phase, WACIE will be carried out by a community institution, and will be a reference organization for WAEMU through the research results and evidence obtained. Through these results, WACIE will be able to feed the process of defining the community directives for WAEMU.

The movement that Benin will instigate through its inclusion in global evaluation networks will lead the countries involved to improve their quality

of governance: this could result in the improvement of their standings according to the Ibrahim Index of African Governance.⁶

CONCLUSION

Development policies and programs at the global level demand better performance and more efficiency not only from states but also from partners and donors. Evaluation is a mechanism used to report on the commitments made in order to ensure good governance: it is a strategic steering tool that can help to inform decision makers.

African states must therefore take ownership of evaluation, according to their sociocultural contexts, and they should pool their experiences in order to improve governance. This pooling calls for the establishment of regional and subregional cooperation platforms for the strengthening of national M&E systems, with the aim of improving the living conditions of the people.

Benin's experience in the field of public policy evaluation has shown that it is important to open up to other experiences in order to capitalize on achievements. It is within this framework that Benin has initiated the setting up of a regional program for capacity building and impact assessment in the WAEMU countries, in partnership with 3ie. This program will not only develop impact assessment within WAEMU countries, but will also strengthen national evaluation systems and the capacities of stakeholders to conduct the evaluation.

To achieve this, WACIE will establish a mechanism for the selection of evaluation areas, and analysis of the results and conclusions of evaluation reports. It will also lead consultations aimed at setting standards, guidelines, and supporting funding mechanisms.

Benin is also relying on the Twende Mbele program, which is the result of its cooperation with South Africa and Uganda, to strengthen the collective learning process at the subregional level through M&E performance improvement and evaluation practice. These cooperative regional initiatives in evaluation will open new perspectives for the development of evaluation in Africa.

REFERENCE

Mendiratta, Vibhuti. 2011. "Report on Impact Evaluation in Sub-Saharan Africa." Working Paper, UMR DIAL, Paris.

⁶The Ibrahim Index of African Governance (IIAG) provides an annual assessment of the quality of governance in every African country. Originally established in association with the John F. Kennedy School of Government at Harvard University, the IIAG now consists of more than 90 indicators in 14 subcategories, 4 categories, and 1 overall measurement of governance performance. These indicators include official data, expert assessments and citizen surveys, provided by more than 30 independent global data institutions. This represents the most comprehensive collection of data on African governance. <http://mo.ibrahim.foundation/iiag/>.