eVALUation Matters
A Quarterly Knowledge Publication on Development Evaluation

Promoting an Evaluation Culture in 2020 and Beyond

First Quarter 2020
eV ALUation Matters

Is a quarterly publication from Independent Development Evaluation at the African Development Bank Group. It provides different perspectives and insights on evaluation and development issues.

Editor-in-Chief:
Mirianaud Oswald Agbadome - Principal Evaluation Officer

Acknowledgments:
IDEV is grateful to all contributors, reviewers, editors, and proofreaders who worked on this issue, in particular: Karen Rot-Münstermann, Dieter Gijbrechts and Tomas Zak.

Editing and translation:
Dieudonné Toukam; Melora Palmer

Design & Layout:
Créon (www.creondesign.net)

Photo credits:
© African Development Bank Group

About Independent Development Evaluation
The mission of Independent Development Evaluation at the AfDB is to enhance the development effectiveness of the institution in its regional member countries through independent and instrumental evaluations and partnerships for sharing knowledge.

Disclaimer:
The views expressed in this publication belong solely to the authors, and not necessarily to the authors' employer, organization or other group or individual they might be affiliated with.
First Quarter 2020

A “culture of evaluation” is often lauded as a key solution to improving the effectiveness and efficiency of governments, development organisations and international financial institutions alike. The extant literature suggests that for results, performance management, evaluation and research functions to be effective, organizations and institutions need a strong culture of evaluation – one where evidence is deliberately sought in order to better design, implement and deliver interventions. This evaluation culture is premised on a drive for continuous improvement that pervades all aspects of work, and a willingness to learn from failures as well as successes.

This edition focuses on new thinking and strategies on the why, what, and how of promoting a culture of evaluation. It draws on a wide variety of case studies from various institutions and African countries to explain why an evaluation culture is important, what it should look like and which tools, systems and processes are important for institutions and governments to promote it. Finally, it also suggests new avenues to widen and deepen the evaluation culture in Africa.
From the Evaluator General’s Desk

Roland Michelitsch, IDEV, African Development Bank

This edition aims to expand the thinking around evaluation culture. It provides a diversity of insights on what evaluation culture means for various organizations, networks and countries, whether and why it must be promoted and finally how to create an inclusive evaluation culture in 2020 and beyond.

Institutional and governance architecture as enablers of an evaluation culture: The experience of IFAD

Oscar A. Garcia, Director, and Fabrizio Felloni, Deputy Director, Independent Office of Evaluation, International Fund for Agricultural Development

Drawing on the experience of IFAD, this article discusses the building of an organizational evaluation culture and how independent evaluation can contribute to it. It argues that an evaluation culture goes beyond the formalities and procedures that may exist in an organization, and that the presence of an independent evaluation function may not generate, on its own, an evaluation culture.

Use of self-evaluation in International Financial Institutions

Rasmus Heltberg, Lead Evaluation Specialist, Independent Evaluation Group, World Bank Group

This article reports and reflects on the author’s experience with a 2016 evaluation of the World Bank Group’s self-evaluation systems. He also provides some suggestions for improving self-evaluations and increasing their use and value for organizations.

Status of Development Evaluation Culture in Ethiopia: A Desk Review

Hailu Filmon, Monitoring and Evaluation Consultant

This article examines Ethiopia as a case study and scores nine different aspects of Ethiopia’s evaluation culture. The author argues that while an evaluation culture in Ethiopia is recognized as important, this commitment has not yet been integrated into the political and administrative system of the state.
Beyond Monitoring and Evaluation Plans: Towards an evaluation culture in Ghanaian Local Government
Clement Mensah, Consultant, IDEV

This article examines the disconnect between Ghana’s national evaluation culture and its evaluation culture at a local level. It suggests that promoting an evaluation culture in local government should go beyond simply designing M&E plans to also focus on creating a conducive enabling environment.

Interview: Reflecting on the evaluation culture at the African Development Bank
Armand Nkoyimana, Manager, Department for Delivery, Performance Management and Results, and Roland Michelitsch, Evaluator General, African Development Bank

Armand and Roland share their insights on how the concept of an evaluation culture has gained momentum over the last few years and what this means for the AfDB as well as for African countries.

A new generation of advocates: young and emerging evaluators and the promotion of a global youth-participatory evaluation culture
Fazeela Hoosen, Director, Benita Williams Evaluation, and Kenza Bennani, Evaluation Manager, Ministry of Foreign Affairs, France

In the last five years, Young and Emerging Evaluators have increasingly contributed to shaping global evaluation culture. In doing so, they have positioned themselves as a new generation of advocates for a “youth-participatory evaluation culture”, one that harnesses the values and skills that youth and YEEs can bring to the evaluation table. In this article, the authors make the case for YEEs to fuel a movement towards the achievement of a global youth-participatory evaluation culture.

Promoting a culture of evaluation in 2020 and beyond

This article discusses what is involved in the establishment of a culture of evaluation and what can be done to create and maintain such a culture in African countries. The article argues that developing a culture of evaluation requires deliberate efforts and strong leadership at a national level, as well as interested stakeholders and a commitment to transparency and good governance.

The culture of evaluation in the age of big data: The need for a new evaluation paradigm for the 4th Industrial Revolution?
Michael Bamberger, Independent Evaluation Consultant

This article discusses the dramatic changes that will take place as the 4th Industrial Revolution evolves and how this will affect the “culture of evaluation” in Africa. It questions how well the transformation will be managed: will evaluation remain a major force in the new ecosystem or will the sector be slow to adapt, such that this great opportunity is lost and evaluation becomes marginalized in the brave new world of the 4IR?

News in pictures

Hot off the Press

Published
March 2020

Editorial Calendar

Q2 2020
Preparing evaluation of the future: Big Data, modern technologies and shifts in global development priorities

Q3 2020
AfDB Evaluation Week 2020

Q4
CSOs and NGOs in Evaluation

Table of contents
In addition to promoting accountability and learning, the third leg of IDEV’s mandate is to promote an evaluation culture within the African Development Bank as well as in its Regional Member Countries.

At IDEV, we understand that achieving this mission requires a multipronged approach. We do this by not only conducting credible evaluations and disseminating our reports, but also by tracking how the Bank responds to our evaluations and implements their recommendations. We have also strived to engage more comprehensively with departments and teams so that they will find our evaluations useful in improving their work.

As a result, we have seen a substantial increase in the use of our evaluations as well as more demand for evaluations. Our ultimate goal is to ensure that the Bank becomes a learning organization that requests evaluations, produces credible evidence and feeds it into a continuous improvement of the organization’s policies, processes and interventions. Recent examples of the demand for and interest in evaluation at the Bank have been the requests for our help in evaluating the Bank’s Implementation of its Development and Business Delivery Model and the Middle Income Countries Technical Assistance Fund. Our evaluations of Program-Based Operations, the Quality Assurance Processes of the Bank, and the Bank’s
Integrated Safeguards System have been good examples of strong engagement with the relevant Bank policy and operations departments, which led to greater buy-in and uptake of evaluation recommendations.

Raising the interest in evaluative knowledge and integrating evaluation into the daily work of the Bank certainly goes a long way to promoting a culture of evaluation. However, creating an evaluation culture at an institution such as the Bank is assuredly a long-term process. Sharing from my experience in another organization, it took about 15 years to fully embed results measurement into corporate and individual incentives— which is essential to change the culture. These comprehensive steps included setting up a solid independent ex-post evaluation system, translating development results into the institution’s key performance indicators, strengthening the monitoring system, establishing a long-term performance award system based on development results achieved, and fueling a system of forward-looking development goals and performance objectives at the institution level, cascaded from top management to operational staff. With that experience, it is my conviction that we can achieve this faster here at the Bank.

Creating an evaluation culture and the system to support it needs continuous innovation along with rigor and adaptation with the goal of improving the development impact of the Bank. To this end, the implementation of our independent evaluation policy remains central to our action while constantly finding ways to strengthen evaluation practice and ensuring the interest of our clients and stakeholders for the evaluations we produce. It is also not an agenda we can pursue alone—we must work together with the management and the governing bodies of the Bank.

In parallel, we are investing a lot of effort in promoting an evaluation culture in African countries. This immense task involves partnering with others on various initiatives. Among others, we have been supporting and engaging with national and continental evaluation networks like the African Evaluation Association (AfrEA), hosting the secretariat of the African Parliamentarians’ Network on Development Evaluation (APNODE), assisting young and emerging evaluators, and working with governments to build national capacities in evaluation through the Twende Mbele peer learning partnership.
Building robust national evaluation systems is a huge undertaking that requires political leadership and national commitment to the development and use of evaluation. In this regard, the African continent has witnessed positive examples and initiatives that are laid out in this edition and that could inspire many countries. The 2030 Agenda for Sustainable Development, with its focus on tracking and reporting on development results through national evaluation systems, represents a real opportunity for many countries to move rapidly and efficiently into promoting their own model and a national evaluation culture that all development partners including the AfDB will be ready to support.

This edition of Evaluation Matters aims to expand the thinking around evaluation culture. It provides a diversity of insights on what evaluation culture means for various organizations, networks and countries, whether and why it must be promoted and finally how to create an inclusive evaluation culture in 2020 and beyond. We have articles on evaluation culture in international financial institutions and in African countries; views from experienced as well as young and emerging evaluators. While some articles are forward looking, they all boil down to the same conclusion: there is no unique model to promote an evaluation culture, however some key elements such as enduring leadership, participation, clear frameworks, tools and resources are necessary.

Promoting an evaluation culture is a constant challenge for all. With this flurry of opinions and perspectives, it is my hope that these contributions will not only inform, but also inspire others to join the collective effort in making evaluation an effective development and learning tool on the African continent.

Happy reading!

About the Evaluator General

Roland Michelitsch is the Evaluator General of the African Development Bank (AfDB). Prior to joining the AfDB, he spent many years with other Multilateral Development Banks. At the Inter-American Development Bank (IDB) Group, he led evaluations of private and public sector activities. With the International Finance Corporation (IFC), he managed the investment unit of the Development Impact Department. He also led IFC’s project evaluation system and framework, and evaluations on various topics. In the World Bank, he worked on corporate governance in Central and Eastern Europe (and in sub-Saharan Africa on population, agriculture and the environment). Roland holds a PhD and MA in Economics from the University of Arizona and an MBA from the University of Graz.

This article discusses the building of an organizational evaluation culture and how independent evaluation can contribute to it, using the case of IFAD as an example. Creating an evaluation culture is distinct from introducing, enforcing and complying with norms, bylaws and processes. It goes beyond the formalities and procedures that may exist in an organization. The article introduces conceptual clarity on organizational culture, organizational evaluation culture, the key conditions and enablers, and how independent evaluation can support their development, before setting out the experience of IFAD.
An organizational culture can be understood as a set of common values, beliefs, goals and business processes, irrespective of their level of formality (Ravasi and Schultz, 2006). Such culture permeates the strategies, style and modus operandi of an organization.

An organizational evaluation culture can be defined (Mayne, 2010) as a culture which appreciates and deliberately demands data and empirical information on its performance and outcomes and uses findings from such evidence in order to improve its performance and the achievement of outcomes. Mayne identifies some of the characteristics of an organization that possesses a strong evaluation culture: (i) engage in self-reflection and self-examination (i.e., self-evaluation); (ii) deliberately seek evidence on what it is achieving; (iii) use results information to challenge and support what it is doing; (iv) value candour, challenge and genuine dialogue; (v) engage in evidence-based learning and make time to learn; (vi) encourage experimentation and change. Such evaluation culture clearly presents advantages in terms of capacity and timeliness to correct the course of action, adaptation to change and adherence to corporate mandate.

A consolidated practice of data gathering can help but, on its own, does not necessarily drive an evaluation culture. An important corollary of the above definitions is that the availability of empirical data and the capacity to analyse them is not a sufficient condition to establish an evaluation culture. Even when data are available, and if data quality or representativeness is not an issue (a heroic assumption in most cases), an evaluative culture would need:

- capacity to analyse and interpret data, including at the higher decision-making levels;
- deliberate interest in findings from data analysis to undertake an objective review of the operational performance and results, connected with some reward system to create and maintain incentives;
- a feedback loop that links findings to action to modify current and future operations and business modalities.

Key Messages

- An organizational evaluation culture is one that demands and uses empirical information to assess performance and outcomes and orient managerial decisions.
- In addition to the availability of data, there needs to be capacity to analyse and interpret them, deliberate interest in findings and a feedback loop leading to action and changes.
- The presence of an independent evaluation function may not generate, on its own, an evaluation culture but represents an important building block.
- Compliance with the evaluation procedures is not equivalent to evaluation culture. It is necessary to move beyond formalism and there is no substitute to the commitment of leadership (Management and Governing Bodies).
In turn, this requires a strong drive from the organizational leadership, specific organizational arrangements and structural support, and a learning focus. As an example, the systematic conduct of impact evaluations, using rigorous methodologies for controlling sample bias, can set the premises to make decisions based on sound evidence. However, these evaluations would have to be intentionally utilised to support managerial decisions. As noted in a World Bank evaluation (IEG 2012), the results of impact evaluations were not used systematically as a basis for decision making, due to the absence of a clear feedback loop and system of incentives. A more recent evaluation by the Inter-American Development Bank (OVE, 2017) found that some impact evaluations were instrumental for the Bank to avoid ineffective projects, although the overall findings on the use of impact evaluations were uneven.

Evaluation in international organizations

In the experience of many international organizations, the presence of a dedicated and independent evaluation function was one of the building blocks to introducing or promoting an evaluation culture. As noted, leadership plays a key role in an evaluation culture. In larger organizations, leadership commitment to evaluation may be prone to the principal-agent problem (Laffont and Martimort, 2002) with the related issues of moral hazard and conflict of interest. For this reason, one of the gold standards of evaluation practice [OECD/DAC, Evaluation Cooperation Group and UN Evaluation Group] is that the responsibility for the evaluation of projects, programmes and policies should not be in the hands of persons or organizational units that were previously responsible for their design and implementation.

In many international financial institutions such as the multilateral development banks, the independent evaluation function is placed under the direct reporting line of the Executive Board. The Executive Board represents (a sub-set of) the membership of the organization (the principal). The fact that the evaluation unit does not functionally report to the management of the organization (the agent) addresses problems of moral hazard and conflict of interest.

However, an important prerequisite for the system to function well is that the Executive Board (principal) continues to value and demand empirical evidence and independent analysis, and keeps management (agent) accountable for results and for implementing recommendations.

IFAD’s experience

Historically, this has been the trajectory followed by the Independent Office of Evaluation (IOE) of the International Fund for Agricultural Development (Figure 1). The office started as a monitoring and evaluation unit (i.e. self-evaluation) in the Department managing development projects, and was then turned into a semi-independent unit, reporting only to the President (chief executive) of the organization. It became fully independent in 2003, under the direct reporting line to the Executive Board.

While independence can help deal with the principal-agent problem, it cannot be solely relied upon to build an evaluation culture in the whole organization. As noted, there needs to be a clear drive from the organizational leadership (both the principal and the agent) and an intentional use of data and information to assess, learn and re-orient programmes, strategies and internal processes. If the mandate to assess organizational activities is confined within an independent unit, there is no guarantee that it will create strong incentives across the organization to seek information on programme performance, set time to learn and innovate. Two practical constraints relate to the following:
Flow and reach of information: in a large organization, the availability of an evaluation report does not automatically translate into shared knowledge on its findings and recommendations in all organizational units.

Timing and timeliness of feedback: Depending on the stage of a programme or strategy within its lifecycle, the type of information and the level of analysis required vary (e.g., on outputs, intermediate outcomes, longer-term impacts). Most independent evaluations take place towards project end-line and tend to be summative (i.e. assessing the results), while implementers at an earlier stage need formative feedback (i.e., reviewing the implementation progress and process and its consistency with design).5

The importance and the role of the self-evaluation function. For this reason, many international development organizations distinguish between an independent and a self-evaluation evaluation function. While the independent evaluation function is placed under the supervision of the Executive Board, the self-evaluation function is a task of the operational units under the ultimate responsibility of the senior management.

Self-evaluation is a very important building block for developing an evaluation culture in an organization. A well-performing evaluation system is based on the premise that assessing performance and results is a duty of managers at all levels of the organization. Self-evaluation can also be an important building block for independent evaluation by providing useful secondary data for independent review and validation and triangulation with additional sources. The cost of generating, ex post, basic information on the implementation performance of a project or programme would be very high. In the words of a former Director General of the Independent Evaluation Group of the World Bank, “self-evaluation is to independent evaluation the equivalent of what accounting is for audit”6.

Independent evaluation can play a key role in raising the standards of self-evaluation. Evaluating a project or programme is not an activity that is naturally familiar to everyone. It requires a certain level of skills, competencies, subject matter technical knowledge, intellectual rigour and mind-set for enquiries. Units that are primarily in charge of designing and supporting the implementation of programmes may not have the expertise, time and focus that are required to conduct high quality self-assessments.
In many international financial institutions, one of the functions of self-evaluation is to assess and report on completed projects. From its side, the independent evaluation function is typically responsible for: (i) helping management establish standards for self-evaluation and build capacity; (ii) assessing the quality of self-evaluation products; (iii) conducting independent validation work at the project level, while devoting most of its resources to higher-level and more strategic evaluation products; and (iv) periodically carrying out an independent assessment of self-evaluation systems in the organization, which can help assess the strength of the evaluation culture.

IOE at IFAD initially devoted most of its effort to evaluating projects. Self-evaluation had a limited role within the organization. The fundamental assumption was that project management teams, under the aegis of borrowing country governments, would establish their own monitoring and evaluation units and provide reporting during implementation and at completion point. At the corporate level, limited priority was given by management to assessing the quality of work of project M&E units and proactively reviewing progress made by projects.

From 2006, the first steps were made between IOE and IFAD management to adopt a common set of criteria for evaluating projects, based on similar definitions, as well as a common ratings system, based on a six-point symmetric scoring (harmonization agreement). This assessment system was based on the evaluation methodology of the independent evaluation function, which was progressively adopted and internalized by the management for its self-evaluation and became part of its business processes.

From 2010, after a peer review of IFAD’s evaluation function, the Executive Board decided that IOE would conduct a validation of all project completion reports (prepared by the governments with IFAD management support), based on desk review of the documentation available (Figure 2). On a more selective basis,
IOE would evaluate a smaller number of projects based on country visits. The three-fold purpose of this decision was to:

- provide the evaluation function with an overview of the quality of self-evaluation activities;
- through systematic scrutiny, generate incentives for Management to improve the quality and credibility of self-evaluation activities; and
- obtain some form of independent assessment on all operations that are completed.

From individual self- and independent evaluation products to consolidated analysis. To provide a more consolidated review of findings, both the self- and the independent evaluation function prepare an annual synthesis report of their respective findings (Figure 3). At IFAD, these are known, respectively, as the Report on IFAD Development Effectiveness or RIDE (self-evaluation reporting) and the Annual Report on Results and Impact of IFAD Operations or ARRI (consolidating findings of independent evaluations). These are both presented to the Executive Board of IFAD and offer the two perspectives on the implementation performance and outcomes of the operations.

The preparation of two reports from the self- and the independent evaluation function is not a duplication but a reinforcement of the necessary interaction (feedback loops) and dialectics between the two functions. While both report at project level, where there is better comparability (thanks to a harmonization agreement on the methodology), the RIDE also reports on other levels of results, which are periodically agreed upon with IFAD member countries. On the other hand, the ARRI also reports on findings from higher-level evaluations (such as country level evaluations, thematic evaluations, corporate-level evaluations and evaluation syntheses).

It is inevitable that self- and independent evaluation will offer different perspectives and analytical angles. The comparison between the two, while potentially contentious, is also a source of knowledge.

**Figure 3: Consolidating self and independent evaluation findings**

![Diagram showing the consolidation of self-evaluation and independent evaluation findings](source: Elaboration by the authors based on IFAD Evaluation Manual (2015))
For instance, self-evaluation is more likely to take into account the constraints encountered during project design and the challenges related to the political environment in which projects operate. On the other hand, independent evaluations are typically more straightforward in assessing the level of achievement of objectives and results and in questioning the assumptions of the given projects.

Tracking responses to evaluations and their recommendation helps signal the urgency of implementing recommendations. An evaluation culture ultimately aims at producing evidence on performance and results in order to drive improvements and change. The conduct of an evaluation, even if of high quality, does not guarantee that recommendation will be heeded. A system that helps track recommendations and transparently records their follow-up can help enhance commitment to recommendation follow-up.

At IFAD, a streamlined process exists to follow up on the recommendations and help close the learning loop (Figure 4). All recommendations of evaluations completed in a given year are tracked by IFAD management (PRISMA report) and, along with the proposed action plan, are presented to the Executive Board in the following year. IOE provides comments on the quality of follow-up, highlighting gaps, if required. Table 1 illustrates the IFAD Management tracking of recommendation follow-up for the period 2010-2018. It highlights that 65 per cent of the recommendations had been fully followed up, 25 per cent had an on-going follow-up and 3 per cent partial follow-up. Outright disagreement had happened in 1 per cent of cases.

| Table 1: Follow-up status of evaluation recommendations at IFAD (2010-2018) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | Full Follow-up | Not applicable* | Not agreed | Not yet due | Ongoing | Partial | Pending | Total |
| Number | 1027 | 39 | 11 | 27 | 402 | 44 | 38 | 1588 |
| % | 65% | 2% | 1% | 2% | 25% | 3% | 2% | 100% |

* Not applicable recommendations relate to project investments that IFAD was initially envisaging but have not been pursued

Source: IFAD (2019)
Reflections

It is necessary to move beyond formalism. The above general features and the specific examples from IFAD relate to a rather mature evaluation environment, where processes and roles are well defined. However, on its own, complying with the processes does not ensure the presence of a strong evaluation culture. To take an example, while management responses may show wide acceptance of evaluation recommendations, this refers to a future course of action and it cannot be taken for granted that the follow-up will happen as presented. In principle, this could be assessed through an independent review exercise looking retrospectively at the quality of evaluation follow-up. So far, there has not been a comprehensive retrospective assessment of the quality of follow-up to evaluations at IFAD.

As the conceptualization of Mayne (2010) illustrates, there is no substitute to the leadership commitment. This refers to the commitment of both the governing bodies (such as the Executive Board) and the management of the institution (i.e., both the principal and the agent) to create an operational culture that is nurtured by evidence on performance and results, where there is some tolerance for risk and errors, and a strong drive to make improvements based on evidence. Usefulness of evaluation ultimately depends on the organizational commitment to using findings for improved decision-making. In an institution like IFAD, assessment, self-reflection and continuous learning are key to address complex and dynamic challenges such as rural development, rural poverty reduction and rural transformation.

The presence of an independent evaluation function may not generate, on its own, an evaluation culture but can contribute to such culture through validation of evidence and analysis, generation of primary evidence, setting standards and conducting periodic evaluations of self-evaluation systems.
Institutional and governance architecture as enablers of an evaluation culture: The experience of IFAD


References

Endnotes

1. The authors are, respectively, Director and Deputy Director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD). The opinions expressed in this article are only those of the authors and do not necessarily reflect those of the organization for which they work. Contact author: f.felloni@ifad.org

2. Adapted from Mayne, 2010

3. Such problem arises when one stakeholder (the agent) is tasked to make decisions and choices representing another stakeholder (the principal), leading to risks of moral hazard and conflicts of interest.


5. An early distinction between formative and summative evaluation was proposed by Scriven, M. (1967).

6. Prof. Robert Picciotto, personal communication.

7. Three scores for positive performance / results (6-highly satisfactory; 5-satisfactory; 4-moderately satisfactory) and three scores for negative performance / results (3-moderately unsatisfactory; 2-unsatisfactory; 1-highly unsatisfactory).
Oscar Garcia is the Director of the Independent Office of Evaluation at the International Fund for Agricultural Development (IFAD). Garcia is the Vice-Chair of the United Nations Evaluation Group, and he was the Chair of the Evaluation Cooperation Group of multilateral development banks in 2017. Before joining IFAD, Garcia served as head of the advisory services at UNEP - Technology, Industry and Economics Division, Paris, providing guidance to the Partnership for Action on Green Economy. Garcia holds a Master’s degree in organizational change management, New School University, New York, and an MBA from the Bolivian Catholic University. He holds a BSc in Economics, University of Santa Catarina, Brazil. He authored several evaluation reports and papers on development evaluation.

Fabrizio Felloni is the Deputy Director of the Independent Office of Evaluation within the International Fund for Agricultural Development (IFAD). He was Lead Evaluation Officer at IFAD from 2010 to 2016. He was an Evaluation Specialist at the United Nations Development Programme from 2008 to 2010 where he prepared their first methodology for country-level evaluations. He holds a Master’s degree in Agricultural Economics from Washington State University (USA) and a Master-equivalent degree in Social and Economic Sciences from Bocconi University (Italy).
Weakness in evaluation culture may stem from weaknesses in the organization's evaluation system. An evaluation of the World Bank Group's self-evaluation systems finds that there is comprehensive coverage but that systems produce little value to staff, line management, clients, and beneficiaries. Information generated through the systems is not regularly mined for knowledge, learning, or performance management by operational units. There is for instance a widespread blaming and frustration, often around ratings disputes, that limits the appetite to learn from evaluations. Managers and evaluators looking to improve evaluation culture and learning from evaluation systems may want to focus on how well mandatory self-evaluation promotes evaluation utility. For example, is there a way to better target evaluations to decision-makers' key challenges and evidence needs?
To improve learning from evaluation systems, focus on how the organization uses and values evaluation

An evaluation unit has existed within the World Bank Group (WBG) since 1970. As such, the World Bank has a long history of evaluation and a strong evaluation culture that has grown in size and importance over the years. While initially focusing on project evaluation, the World Bank started self-evaluation of its investments in 1975 and, over time, systems have grown increasingly comprehensive in scope and coverage. Today, self-evaluation lies at the heart of how the WBG and other International Financial Institutions (IFIs) measure their results.

Self-evaluation is done by or for operational units; it builds on monitoring data and feeds into independent evaluation and corporate results measurement, usually via ratings that are independently validated by the independent evaluation offices of the IFIs.

This article reports and reflects on the author’s experience with evaluating the WBG’s self-evaluation systems (World Bank 2016). That evaluation sparked efforts to improve monitoring and reform the WBG’s self-evaluation systems. It has also attracted interest from other IFIs, many of which have self-evaluation systems with similar characteristics and, arguably, pathologies.

Expanding the evaluation framework

Past assessments of self-evaluation systems in IFIs often focused on the degree to which actual practice complied with stated policy and with implicit notions of what an ideal system looked like. Behavioral aspects received limited attention. These assessments often recommended expansion of coverage, assuming that would be a good thing in itself. Assessments didn’t always examine the benefits of doing more evaluation.

The framework for IEG’s evaluation of the WBG’s self-evaluation systems expanded on past assessment frameworks in some ways. It looked at the entire chain from project monitoring and supervision to completion reporting. “Validation”
Use of self-evaluation in International Financial Institutions

- Evaluation Matters First Quarter 2020

It focused on mandated self-evaluations but also considered impact evaluation as an example of voluntary self-evaluation.

It analyzed evaluation systems at multiple levels: templates and guidelines; business processes, data streams, roles and responsibilities; and behaviors influenced by norms, values, incentives, and organizational culture. The evaluation also examined how independent validation and evaluation influence self-evaluation processes, behaviors, and incentives.

**Limited use of self-evaluation information**

The stated purposes of self-evaluation give equal weight to accountability, project/program performance management, and organizational learning. But in practice the operation of systems has focused on results reporting for accountability purposes. The systems produce corporate results measures such as the percentage of operations rated satisfactory.

These systems produce little value to staff, line management, clients, and beneficiaries. There has been less attention to developing learning to enhance performance. Information generated through the systems is not regularly mined for knowledge and learning except by IEG. Project and portfolio performance management also doesn't much rely on evidence from self-evaluation systems. The evaluators encountered widespread blaming and frustration.

We produced a ‘systems map’ to understand the World Bank Group’s evaluation culture and the role of self-evaluation in it. We found that we needed to map the organizational culture in order to understand deeply how actors engage with the evaluation systems and why there was so little learning from them. Figure 1 illustrates some ways in which behaviors, incentives, and organizational culture limit the systems’ utility. There is much focus on project outcome ratings, exacerbated by a competitive organizational culture. Focus on avoiding negative ratings can limit candor. Depth and accuracy in how self-evaluations are written and used suffer. Operational managers face a relentless pressure to keep the pipeline of new projects and disbursements flowing. They have few incentives to invest in project monitoring. Without good data, self-evaluation systems create little value for operational units.

This systems map helps us understand why the perceived value of the knowledge created from self-evaluation systems is low, that too many risks and failures are hidden, safe spaces to learn from failure are missing, and lessons and knowledge are not mined.

The problem is with the mandatory systems, not staff. Staff are genuinely interested in evaluative findings they consider valuable. For example, the World Bank’s impact evaluations tend to be respected and used by staff. Impact evaluations are based on data and methods perceived to be reliable and are carried out in a process that is more congruent with organizational culture. Staff therefore value them more than they do most mandatory self-evaluations.

**Ways forward for more use of self-evaluation**

The issues discussed have been noted in a few individual evaluations and studies (e.g., Weaver 2010; Raimondo 2015; 2016). Reviews of DFID noted heavy reporting burdens on staff and partners that do not seem justified given the use made of the results data, as well as shifting of focus and of incentives toward individual outputs, activities.
Use of self-evaluation in International Financial Institutions

and short-term results (OECD DAC 2014; ICAI 2015). But development agencies have by and large not done much to reform how their cultures promote evaluation use.

Arguably, the move toward evaluation systems that are comprehensive, transparent, and mandatory has been fueled by a belief that evaluation is a universal good (that is, a good idea in all circumstances). This belief has led to rules-based, comprehensive evaluation systems. IFIs tend to run their self-evaluation systems using rules-based approaches that spell out what, when, and how operations should be evaluated. Rules that apply across the board help reduce complexity and make things more manageable in large public sector bureaucracies. But rules that apply across the board may also have undesirable unintended consequences such as changes in behavior, large hidden costs, and avoidance of risk and blame (Dahler-Larsen, 2012; Hojlund, 2014; Raimondo, 2016).

What might be a way forward? For starters, consider what evidence decision-makers really need. Is universal reporting based on across-the-board
evaluation coverage needed? Is there a way in which evaluation could be better targeted to decision-makers’ key challenges and evidence needs? Dahler-Larsen (2012) advocates for a more selective approach to evaluation whereby operations are evaluated when the expected marginal benefit of evaluation exceeds the cost. He proposes a revival of the evaluability assessment:

In other words, to foster a more use-focused evaluation culture, focus on producing evaluations and self-evaluations that will be used and valued by the organization. This would be a marked departure from current practice among IFIs.

References


Rasmus Heltberg is a Lead Evaluation Specialist at the Independent Evaluation Group (IEG) of the World Bank Group. He leads large corporate evaluations, specializing in organizational effectiveness. In his 17 years with the World Bank he has also worked on social protection, social development, adaptation to climate change, and disaster response. He was a member of the core team for the 2014 World Development Report on Risk and Opportunity: Managing Risk for Development. His research has been published in a variety of interdisciplinary development journals, including World Development, Journal of Development Studies, Environment and Development Economics, and Global Environmental Change. Rasmus holds a Ph.D. and M.Sc. in Development Economics from the University of Copenhagen and a M.Sc. in Quantitative Development Economics from the University of Warwick.
Ethiopia has been experiencing a growing demand for good governance and effective decision-making, necessitating evaluation in different sectors of public management. This article aims to assess the status of development evaluation culture in Ethiopia using the approach of Furubo et al. (2002). The approach applies nine indicators of evaluation culture, giving a score to each after verifying the existence of empirical materials and assessing how well the country is doing according to the indicators. The score for each indicator is usually given through consultation between the rated entity and national researchers. However, for this article, I have rated the indicators by reviewing available research on Ethiopia. The findings of the research show a mixed result: Ethiopia has a low score with regard to the use of evaluation in many policy domains; a medium score on Degree of Institutionalization – Government; and a high score in the case of existence of professional organizations. From this we can conclude that an evaluation culture is not yet fully established in the country. On the basis of the analysis, I make some recommendations to strengthen the evaluation culture in Ethiopia.
Hailu Filmon, Monitoring and Evaluation Consultant

Key Messages

- This article assess the status of development evaluation culture in Ethiopia by scoring nine indicators of evaluation culture.
- The overall result of the research implies that Ethiopia does not have a mature evaluation culture as all nine criteria are not satisfied.
- It recommends that efforts to establish an evaluation culture should be integrated into the political and administrative system of the public sector and that government should work with stakeholders in strengthening evaluation culture.
- adopting an evaluation policy should be supported by an enabling environment to improve evaluation culture within a state.

Introduction

Development interventions need to be monitored and evaluated in order to show the impact of actual results, gain lessons from their implementation, and improve future design. This becomes practical when a proper analysis tool is used to assess changes attributable to the specific intervention in a given area or on beneficiaries. Among these tools, development evaluation is considered a prominent one.

Development evaluation is widely understood as the evaluation of development interventions to include projects, programs and policies. Patton (2012) defines development evaluation as a generic term for evaluations conducted in developing countries, usually focused on the effectiveness of international aid programs and agencies.

As a tool of analysis, evaluation has evolved and changed as development issues grow. Development evaluation has been practiced for more than 72 years in the developed world since the inception of the World Bank in 1944 (Imas and Rist, 2008). But, its application in developing countries for their own development interventions, tells a different story.

Ethiopia has been practicing evaluation (Musumba, 2018). However, the country has limitations in applying development evaluation extensively and in a structured manner. As a result, only a few studies have been undertaken so far that reveal the status of development evaluation and degree to which an evaluation culture has taken root in the country.

Materials and methods

Different institutions and researchers have tried to assess the status of development evaluation and evaluation culture within countries. Among the available research, this article follows the approach of Furubo et al. (2002), which helps to identify forces that affect the application of evaluation in different national contexts and the consequences these have
in promoting its use. The approach includes the following:

- Historical overview of supply and demand in relation to evaluation; shifts in institutional settings;
- Description of professionalization in the field of evaluation;
- Description of evaluation in the executive branch of government;
- Description of evaluation in the legislative structure;
- Concluding observations on where national systems are at present and where they may be going; and
- Investigate possible explanations for changes and trends, and implication of evaluation to national developments.

Furubo et al. (2002) also examined nine variables of evaluation culture, giving each a score between 0 (low), 1 (medium) and 2 (high), with all needing to be fulfilled before an evaluation culture can be completely established. Table 1 lists the variables.

**Assessment**

I assessed Ethiopia on these criteria based on information collected from Alemu and Latib (2013) and the Independent Evaluation Office of the UNDP (2015), by making analysis and observations on Furubo et al.’s (2002) nine assessment criteria of evaluation culture. As a result, I have rated Ethiopia’s performance on the indicators as shown in Table 2.

**Evaluation takes place in many policy domains**

This score of 0 for this first criterion implies that evaluation activities in Ethiopia remain a separate function and are only concentrated in few sectors. There are no specific branch of government in charge of a national mandate of the evaluation function which could explained the lack or limited appetite for evaluation studies. Other possible reasons could include a lack of knowledge of evaluation and its contribution towards effective and evidence-based policy making.

**Supply of domestic evaluators in different disciplines**

The country has performed badly in this case. This is supported by the fact that there are currently almost no higher academic learning institutions which offer a curriculum for monitoring and evaluation (M&E) professionals in the country in various fields of study. This is expected, as the demand for evaluation is low at the policy making structure.

<table>
<thead>
<tr>
<th>Table 1: Variables that assess evaluation culture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> Evaluation takes place in many policy domains.</td>
</tr>
<tr>
<td><strong>2</strong> Supply of evaluators specializing in different disciplines.</td>
</tr>
<tr>
<td><strong>3</strong> National discourse concerning evaluation.</td>
</tr>
<tr>
<td><strong>4</strong> Existence of an evaluation sector that remains up-to-date on the norms and ethics of the profession.</td>
</tr>
<tr>
<td><strong>5</strong> Institutional policies in the government for conducting evaluations and disseminating results to decision makers.</td>
</tr>
<tr>
<td><strong>6</strong> Institutional policies are present in parliament for conducting evaluations and disseminating findings to decision makers.</td>
</tr>
<tr>
<td><strong>7</strong> Existence of different stakeholders commissioning and performing evaluations.</td>
</tr>
<tr>
<td><strong>8</strong> Evaluation activities within the Supreme Audit Institution (SAI).</td>
</tr>
<tr>
<td><strong>9</strong> Public sector evaluations that have program or policy outcomes.</td>
</tr>
</tbody>
</table>
As a result, it is hard to imagine and see that due attention is given, and resources allocated, to improve the supply of evaluators in Ethiopia.

**National discourse concerning evaluation**

The country scored the highest in this regard. The Ethiopian government is gaining experience from its implementation of public sector reforms adopted under results-based management to serve as the national strategy (Alemu and Latib, 2013). Adopting results-based management as a public policy management tool will have many implications for the demand and supply of evaluation.

Making government institutions more results-oriented would improve service delivery, the provision of public goods and the use of budget. This will also have implications in terms of accountability, giving space for learning in the process of improving the national governance of policies.

**Professional associations**

In this regard the country is doing well. The available evidence reveals that two professional associations operate in Ethiopia. These are the well-known Ethiopian Evaluation Association (EEvA) and a young professional society referred to as the Ethiopian Monitoring and Evaluation Association, established in 2009 and 2014 respectively.

EEvA has been vocal and active in creating networks with fellow societies in Africa as well as organizing professional conferences (EEvA, Annual Performance Reports 2013–2016a). It is also making efforts to introduce evaluation standards to develop the profession. Although it has done much in terms of increasing its visibility, EEvA’s contribution towards improving the demand and supply of evaluation in the country is quite minimal.

**Degree of Institutionalization – Government**

The score on this indicator is of medium value. The implication being, that even if there are efforts like the support provided by the AFDB through IDEV which strengthened the national M&E system from 2014 to 2018, the country still has limitations in creating the necessary permanent structures to facilitate the commissioning of evaluation studies within government offices and to disseminate findings.

Looking at the historical overview of demand and supply of evaluation in Ethiopia, it seems the country has some experience in the field of development evaluation. Research reveals ▶

### Table 2: Evaluation Culture Ratings in Ethiopia

<table>
<thead>
<tr>
<th>No</th>
<th>Variables</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evaluation takes place in many policy domains</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Supply of domestic evaluators in different disciplines</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>National discourse concerning evaluation</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Professional organizations</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Degree of institutionalization – Government</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Degree of institutionalization – Parliament</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Pluralism of institutions performing evaluations within each policy domain</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Evaluation within the Supreme Audit Institution</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Proportion of outcome evaluations in relation to output and process evaluations</td>
<td>1</td>
</tr>
</tbody>
</table>
that monitoring had special focus especially within the three five-year plans (1957-1973 European Calendar) under the Imperial regime. The available evidence further demonstrates that M&E of country development plans has been intensive in the case of public sector governance, especially during the military regime. However, M&E activities had less focus on impact evaluation during the Derg era (1974-1991 European Calendar). Since the change of the government in 1991, the country articulated an M&E framework which is described in its national development plan. Nevertheless, the focus on monitoring has not been changed over the years resulting in a minimal use of evaluation and evidence-based feedback systems (Alemu and Latib, 2013).

Recently, the Ethiopian Government established the National Planning Council and National Planning Commission by official decree (Regulation No. 281/2013). It has also issued a National Growth and Transformation Plan Implementation Council of Ministers Monitoring and Evaluation Guideline since January 09, 2018. In addition, delivery units have been established starting from the Office of the Prime Minister. These moves demonstrate the significant attention given to establishing an institutional structure that coordinates national level M&E activities and the adoption of a policy framework.

**Degree of Institutionalization – Parliament**

This indicator has a low score. It indicates that currently there are no permanent practices which are responsible within the legislative structure that facilitate the conduct of evaluation studies and dissemination of findings.

Members of the Ethiopian Parliament signed the Yaoundé Declaration of African Parliamentarians on Evaluation in 2014 (Rosenstein, 2015). Nonetheless, it is surprising that the Parliament has not exercised its power to call for any evaluation on government affairs up to now.

The Parliament has a big role to play in establishing evidence-based decision making since it serves as watchdog of executive activities. It also could have impact in terms of spreading pluralism and openness to new ideas in improving the country’s evaluation culture by exercising its mandate. This has greatly affected the demand and supply for evaluation and its contribution in sustaining good governance in the country.

**Pluralism of institutions or evaluators performing evaluations within each policy domain**

The score on this indicator is 1. There are institutions which perform evaluations in Ethiopia; however, they are few. As the demand for evaluation is not existent in the policy domains of the country, it is clear there are not many firms and consultants specialized in the field. This is reflected by market rules.

**Evaluation within the Supreme Audit Institution**

The result for this indicator is zero. This means evaluation activity does not exist within the Office of the Federal Auditor General and it is not currently contributing to the creation of an evaluation culture in the country. Evaluation and audit are directly related. This is because both have the objective of establishing accountability in managing resources and promoting learning gained from the conduct of assessments.

There is a great interest by the parliament for annual reporting on the performance of government institutions by the Office of the Federal Auditor General, to the use of public resources; this has also attracted great media and public attention. However, these federal auditor reports could have been reinforced with more meaningful
impact, in terms of demonstrating the value added to the target group if evaluation studies were integrated within the audit office activities. There is a huge gap in this regard.

Proportion of outcome evaluations in relation to output and process evaluations

The indicator in this regard reveals a score of 1. The implication being that in Ethiopia, many outcome evaluations are not being carried out to show whether interventions had an impact on the problem they were intended to solve. Very few interventions perform outcome evaluations if they are attached to donor interest and donor support is involved. This is because there is a huge knowledge gap about what evaluation is and its contributions to effective decision making in the public sector.

Conclusion and Recommendations

Based on Furubo et al. (2002), this article has defined maturity in two aspects while analyzing the indicators of evaluation culture. The first aspect concerns the degree to which evaluation praxis is established and becomes an integrated part of the political and administrative system. The second feature focuses on the spread and pluralism of an evaluative culture and its openness to new ideas and impulses.

Research shows evaluation is directly related to open public discussion and the concept that policies are open to question. As a result, the Ethiopian government should encourage open public discussion and accustom itself to critical voices through creating opportunities for civil society intervention, including voluntary professional associations in evaluation, by establishing platforms and engaging them on a permanent basis. This is crucial to spreading pluralism and establishing an evaluation culture in Ethiopia characterized by openness and new ideas.

The overall result of the research implies that Ethiopia does not have a mature evaluation culture as there are gaps in certain areas, thus not satisfying all nine criteria. It seems there is recognition for the need of M&E as the institutional structure has been established, however, this commitment is not implemented to the level that it is integrated into the political and administrative system which makes the system functional and effective. Even if there are stakeholders who advocate for the use of evaluation, they are not influential enough to encourage pluralism and exchange of ideas.

"Ethiopia does not have a mature evaluation culture as there are gaps in certain areas, thus not satisfying all nine criteria."

What recommendations can we draw from the assessment to strengthen the evaluation culture in Ethiopia?

The first priority area to improve the evaluation culture relates to the adoption of a policy framework. In this regard, Ethiopia recently issued a National Growth and Transformation Plan Implementation Monitoring and Evaluation Guideline through Council of Ministers on January 09, 2018 (FDRE, 2018).

The second focus area that should be worked on concerns designing programs to implement the policy directions which are stated in the national M&E guideline. In this case, Canada could be a good example, as it introduced the Results for Canadians Initiative whereby it generated demand for evaluation at the national level (Stolyarenko (c)). This is a government management agenda whose focus lies on improving delivery of results, responsible spending, greater transparency and accountability for results.
The third intervention that should be made is to identify evaluation champions. It would be good if these are situated with the government since it plays a leading role. Here, a good example could be Uganda where the President became the lead champion with the government for advocating the use of evaluation in the management of the public sector (Stolyarenko (a)). He also assigned resident district commissioners at the local level to serve as M&E champions. In addition, in Uganda the Office of the Prime Minister is the lead champion institutionally for M&E in the country.

The fourth instrument that should be taken on board is to establish a platform to create public confidence for openness and accountability among stakeholders. In this instance, the case of Colombia Lider, an independent platform of civil society partnership where representatives of the media, banks, foundations, and other organizations are gathered to discuss good governance and social control could be a good example (Stolyarenko (b)). It is used to highlight good performance of municipal mayors, and monitor government plans and spending, and analyze their impact on poverty.

Taken together, implementing these recommendations should improve Ethiopia’s performance on the indicators assessing the strength of evaluation culture in the country.

References


MoFED and UNDP (2012a) Draft Strategy Paper for Mainstreaming of Results Based Management for the Realization of the GTP and Goals in Selected Regions, Addis Ababa, MoFED and UNDP.


Stolyarenko (b), K. Case - Study # 4: National Evaluation Policy in Colombia, Parliamentarians Forum on Development Evaluation.

Hailu Filmon has a Master of Science in Development Evaluation and Management from the University of Antwerp. He has more than seven years of experience in preparing strategic plans, program budgets, project annual work plans, and monitoring and evaluation plans. He has developed and appraised social sector development project technical proposals in numerous government offices and consultancy firms, in different positions. He is a vice president of the Ethiopian Evaluation Association and reference group member of the African Evidence Network.
Effective local governance is an essential precondition for quality public service delivery and better development outcomes. In the same vein, a culture of evaluation is equally essential for local government authorities to execute their development planning and administrative functions effectively, efficiently and sustainably while promoting accountability and learning. That said, a culture of evaluation rarely exists within Ghana’s local government system. Several factors, including limited financial and technical capacities, continue to hinder the promotion of a culture of evaluation amongst metropolitan, municipal and district assemblies. Thus, for Ghana’s local governance system, just designing monitoring and evaluation plans for District Medium-Term Development Plans is not enough. Well-resourced evaluation units coupled with a strong partnership and institutional orientation as well as platforms of exchange are crucial to building a culture of evaluation. This is even more crucial in a time when development partners including multilateral development banks are looking to engage more sub-nationally.
Beyond Monitoring and Evaluation Plans: Towards an evaluation culture in Ghanaian Local Government

Clement Mensah, Consultant, IDEV

Key Messages

- Promoting a strong evaluation culture is as important for local government authorities as it is for central governments, and win-win for national development priorities as well as the Sustainable Development Goals.

- There appears to be little evidence of a widespread evaluative culture within the Ghanaian local government. At best, what exists could be likened to a monitoring culture.

- Promoting an evaluation culture amongst local governments goes beyond just designing monitoring and evaluation plans. The enabling environment for implementing these plans – including financial and technical capacities as well as the right institutional frameworks – should be created.

- Periodic performance assessment tools can be useful in promoting an evaluation culture at the local government level on two fronts. First, they can help highlight key governance lapses and needs, including monitoring and evaluation, which local governments will have to leverage to improve their delivery. Second, such tools have the potential to create greater demand for accountability from civil society.

- National evaluation associations, research institutions and civil society organizations have an important role to play in instilling a culture of evaluation.

Introduction

With the rise in demand for greater accountability and the quest for evidenced-based policies, governments across the world are stepping up efforts regarding monitoring and evaluation (M&E) of policies, programmes and projects. In Africa, M&E practice has undergone significant changes since the turn of the millennium, with a gradual shift towards greater country-owned M&E systems (Basheka, 2015; Mofolo, Mkuyane & Skade 2014). Today, countries like South Africa, Benin and Uganda have relatively advanced national M&E systems, and they continue to exchange and learn from each other (Goldman et al., 2018). Others, like Botswana and Ethiopia, have developed national policy frameworks to guide public sector monitoring and evaluation. The Government of Ghana, following in the footsteps of countries like Benin, established a ministry solely dedicated to monitoring and evaluation of government interventions in 2017, as did Gabon two years later.

This growing prominence of M&E practice in Africa’s public sector stems from not only the need to assure accountability and learning but also due to the “...political recognition of the utility of evaluation to good governance” (Basheka, 2015:77). In other words, there is an intrinsic relationship between M&E and good governance, be it at the national or local level. That said, the evaluation component in M&E is still lagging in practice when it comes to the African public sector (Goldman et al., 2018). Even at the local government level, a similar phenomenon exists. Although M&E practices are becoming popular...
FULFUSO-SAWLA ROAD PROJECT

JONOKOPONTO PRIMARY SCHOOL

INTERVENTIONS:
School Building
Teacher’s Bungalows
Water Facilities
Sanitation Facilities
Sola Power
School Furniture

FUNDED BY:
AFRICAN DEVELOPMENT BANK GROUP
GROUPE DE LA BANQUE AFRICAINE
DE DEVELOPPEMENT

YOUR PARTNER IN DEVELOPING AFRICA
amongst local governments across the continent, subnational M&E systems and plans turn out to be all about monitoring and no or little evaluation, a situation that signals a weak evaluation culture at the local level.

In Ghana, while nearly all Metropolitan, Municipal and District Assemblies (MMDAs) prepare M&E plans to accompany their Medium-Term Development Plans as enshrined in the National M&E Manual and District M&E Guidelines, the component of evaluation is rarely practiced. Meanwhile, like monitoring, evaluation is valuable for accountability and evidence-based policy making (Mackey, 2008).

This article therefore draws mainly on existing literature including M&E plans developed for MMDAs’ as well as annual progress reports to assess the state of evaluation culture in Ghana’s local government system. The article also examines current efforts and tools aimed at promoting a culture of evaluation and proffers recommendations towards an improved evaluation culture amongst Ghana’s MMDAs.

Ghana’s local government system and the role of M&E

Admittedly, the local government system has grown to become an important structure in the governance of development globally. By moving away from a state-centric development administration to a local-level system of development planning and administration, local governance has the power to fast-track a country’s development as it guarantees ownership, leverages local resources and improves local participation. To this end, local governance can be viewed as indispensable to society’s development (Institute of Local Government Studies & Friedrich Ebert Stiftung, 2016). In essence, local governments play a critical role in not just public service delivery but are also seen as vehicles for bringing development to people.

Over the past three decades, Ghana’s local governance system has evolved following the 1988 legislative reforms and was subsequently reinforced by provisions such as the Local Government Act of 1993 and the Local Government ACT, 2010, (ACT 462). As of 2019, there are a total of 260 administrative districts. These districts are the prime administrative, planning, development, budgeting, and rating authorities in their areas of jurisdiction.

In performing these functions, MMDAs, working through District Planning and Coordination Units, develop District Medium Term Plans that are normally four years long and cover a wide range of interventions including construction of school blocks, health centers, and market centers. In doing this, the District Planning and Coordination Units work with other decentralized agencies such as the Ghana Education Service, the Ghana Health Service and the Ministry of Food and Agriculture, amongst others.

The enormity and complexity of the planning, execution, coordination as well as administrative functions of local governments requires the establishment of a system to consistently monitor and measure these efforts and promote organizational learning. It is for this reason that there is a growing need to strengthen M&E capacities at the local level amongst governments, donors and development partners alike (Le Bay & Loquai, 2008).

According to the National M&E Manual, “M&E is the main instrument for assessing the extent to which the government has done what it pledged to do in its national policy statements and development plans.” At the district level, the concept of M&E is becoming popular. Today, all MMDAs are required to prepare M&E plans for their respective District Medium-Term Development...
Plans as stipulated in the National M&E Manual. This plan, which is directly developed by the District Planning and Coordination Unit under the leadership of the District Coordinating Director, has as one of its objectives to capture lessons from the implementation of plans, projects and programmes (see Box 1).

Similar to central government, Ghana’s MMDAs also adopt M&E as part of their governance processes. However, this adoption should not stop with development of plans and the organization of quarterly performance review meetings. Realising the full benefits of such plans while improving development outcomes at the local level will require a strong culture of evaluation.

The State of Evaluation Culture at the MMDA Level

Even though evaluating the effectiveness of government interventions yields several benefits, the absence of a strong evaluation culture – commitment to initiating and utilizing findings of evaluations – is likely to thwart these benefits. Thus, evaluation culture is a critical precondition for evaluation use and utilization (Sandison, 2006).

Within Ghana’s local government system, MMDAs’ evaluation culture can be characterized via multiple lenses. Specific to this review, four main strands are explored.

Legislative and policy environment: At the heart of the quest to promoting a culture of evaluation amongst MMDAs lies the need to create the enabling environment in terms policy and institutional arrangements. While the National M&E Manual provides overall direction for developing sector and MMDA M&E plans, District M&E Guidelines – which are developed and periodically revised by the National Development Planning Commission – serve as the main reference document for the preparation of MMDAs’ M&E plans. The 2014-2017 District M&E Guidelines, for example, detail the institutional arrangements for decentralized M&E, present the sequence of activities for preparing the plan and offer modalities for preparing the plans in a participatory manner.

Within the same guidelines, a specific section is dedicated to development evaluation, where the role of evaluation, its core values, the different types as well as the steps to be followed in conducting one are tersely presented. This information also includes a sample

---

**BOX 1. RATIONALE FOR DEVELOPING M&E PLANS AT THE MMDA LEVEL**

- Assess whether the DMTDP targets were being met
- Identify achievements, constraints and failures so that improvements can be made to the DMTDP and project designs to achieve better impact
- Identify the appropriate interventions that will increase the responsiveness of targeted beneficiaries
- Provide information for effective coordination of district development at the regional level
- Document lessons learned from the implementation of programmes and projects
- Improve service delivery and influence allocation of resources in the districts
- Demonstrate results to stakeholders as part of accountability and transparency
- Reinforce ownership of the DMTDP and build M&E capacity within each District.

Source: Guidelines for the Preparation of the District Monitoring and Evaluation Plan Under
evaluation matrix, which is supposed to come with the District M&E Plan.

While these provisions are important, their practical application remains a significant challenge. A cursory glance at these District M&E Plans and some annual progress reports reveals a large gap in terms of how evaluations are captured. Except in a few instances where mention is made of mid-term evaluation of the DMTDPs, most of the M&E activities captured in these plans are mainly monitoring in nature. This bias partly explains the limited nature of the evaluation provisions in the guidelines and how this needs to change, not just at MMDA level but even at the national level. Currently, Ghana’s national evaluation system appears weak and this obviously cascades into the kind of evaluation culture that exists at the MMDA level.

Technical and financial capacity: Technical and financial capacities are fundamental for effective evaluations (Adrien & Denis, 2008). Following this, promoting and sustaining a culture of evaluation at the MMDA level would require that the necessary technical and financial capacities are in place. For many MMDAs, this is a major hurdle to promoting a culture of evaluation. Evaluation capacity amongst Ghana’s Ministries, Departments, Agencies as well as MMDAs remain limited (Twende Mbele, 2018; UNDP, 2015; Dery, 2013). Indeed, these challenges are clearly enumerated in some M&E plans themselves.

In particular, the context in which the term ‘evaluation’ is used in most M&E plans and annual progress reports is in itself problematic and at times confusing. A review at some of these documents reveals that the concepts of ‘monitoring’ and ‘evaluation’ are used almost interchangeably. For example, in some M&E plans, the term ‘evaluation’ has been used in the context of undertaking activities such as field visits. Even in situations where an evaluation matrix is presented in the plan, the wrong methodologies have been indicated, including site visits and meetings. This further adds to the point on the mixed understanding and use of the two terms. It also points to a pattern of non-compliance owing to limited technical capacity (Dery, 2013).

Utilization: This is an indispensable part of the evaluation cycle and certainly critical to a sustained evaluation culture. Sadly, given that real time evaluation of interventions is rarely conducted at the MMDA level except for mid-term reviews of the DMTDPs, little evidence exists to validate whether both monitoring and evaluation findings feed back into improving local governance practices and development planning processes. Even for emerging performance assessment tools such as the District League Table where dissemination workshops are organized to share findings with key stakeholders, the decision of MMDAs to take up issues identified from the assessment lies with them.

Participation and partnerships: Recognizing the value of evaluation and the consequent demand that comes with it is equally fundamental to a strong evaluation culture. For governments, one effective way of achieving this is by encouraging greater participation and building strong partnerships with stakeholders. At the MMDA level, this can happen in two ways: in the process leading up to generating the M&E plans and during the actual implementation of the plans including dissemination of results.

At the design phase, M&E plans report conducting stakeholder analysis as part of their engagement processes as stipulated in the guidelines. These stakeholders range from non-governmental organizations to community-based organizations. In addition, the M&E guidelines encourage the use of participatory M&E tools such as Participatory Rural Appraisal, Citizen Report Card, Community Score Card, and Participatory Expenditure Tracking Surveys.
Another important area that underpins the participation and partnership strand stems from partnerships with civil society organizations in promoting a culture of evaluation. Already, some of these collaborative efforts are taking place – for example, the Ghana M&E Forum is working with Twende Mbele and CLEAR-AA to strengthen government-civil society organization collaboration in M&E for the sanitation sector – but they normally tend to be nationally oriented rather than local. Obviously a good example that can be explored at the MMDA level.

The role of tools in promoting a culture of evaluation amongst MMDAs

Besides the traditional M&E guidelines and their accompanying plans that are meant to mainstream an M&E culture amongst MMDAs, there a couple of performance assessment tools that have the potential to complement MMDAs in reflecting on their administrative and planning processes and assure good governance and accountability in the long run. Two of these tools are discussed below.

District Assemblies Performance Assessment Tool: This is a tool which serves two main purposes: first, to assess the performance of MMDAs based on a wide range of indicators and second, to serve as a basis appropriating the allocation of District Assemblies Common Fund responsiveness factor grant to MMDAs (Government of Ghana, 2018). Management, coordination and M&E forms one of the ten main indicators used for computing this assessment. While an important tool for incentivizing MMDAs to do better as well as provide general performance feedback, it cannot replace a functional evaluation system. For example, the kind of issues the assessment captures in the area of Management, Coordination and M &E is too limited to provide meaningful lessons regarding project- and programme-specific implementation issues concerning what works and what does not.

Ghana’s District League Table: Since 2014, the District League Table – jointly designed by the Center for Democratic Development and UNICEF Ghana in collaboration with the Ministry of Local Government and Rural Development – has served as an important social accountability tool for measuring how the country’s MMDAs are faring in terms of development and delivery of basic services on an annual basis. Its main aim is to promote social accountability and can be used for programme planning (Center for Democratic Development & UNICEF, 2015). The assessment is based on six key sectors: health, education, sanitation, water, security and governance. Thus, it has the potential to create demand for monitoring and evaluation by civil society in particular. However, this tool does not adequately highlight some of the governance lapses at the local level that ought to be taken up especially in the areas of M&E. In a few cases, monitoring and supervision have been flagged in the District League Table reports.

Instilling an evaluation culture in Ghana’s Local Government System

Despite the fact that the development of M&E plans is a critical feature within Ghana’s local governance system, it tends to be biased towards monitoring. Simply put, what really exists is a culture of monitoring and not evaluation. As this article has shown, the evaluation ecosystem at the local level appears weak. Thus, specific to Ghana, the following recommendations are essential in promoting a culture of evaluation amongst MMDAs.

- Reorient M&E systems to adequately feature evaluation: A first step to building a culture of evaluation
within Ghana’s local government is to reorient the M&E system to adequately feature evaluations. This will entail putting in place a fully entrenched national evaluation system that cascades down to the district level. Also, intensifying evaluation education including building staff capacity to conduct evaluations while ensuring acceptable levels of rigor, credibility and independence is critical. Moreover, the templates used for developing these district M&E plans will need some improvements to appropriately capture evaluations. This will require the joint collaboration of the ministries of M&E and Local Government and Rural Development together with the National Development Planning Commission as well as development partners.

Establish results sharing platforms:
Having a dedicated platform for showcasing experiences in relation to M&E is important. In fact, the M&E Ministry launched the maiden Ghana Results Fair in January 2020. This saw various Ministries, Departments and Agencies showcasing their results. Such a platform can target MMDAs too, whereby MMDAs in respective regions get the opportunity to share their results. Having such a platform brings with it multiple benefits. First, it is an opportunity for peer-learning amongst district assemblies on their M&E practices and engagements. Secondly, it could be a useful tool for disseminating information to their constituents, which in turn will encourage them to demand more results.

Strengthen partnerships with research centres and national evaluation associations: Given that national evaluation associations are a vital force in promoting evaluation culture, strengthening partnerships with such actors including the Ghana M&E Forum is an important starting point. Since budgets are a major constraint to creating functional M&E within the public sector, partnering with research institutions and national evaluation associations to support evaluation capacity building initiatives, exchange programmes as well as joint evaluations becomes imperative.

Develop evaluation-specific accountability and learning tools:
As mentioned earlier, tools such as the District League Table are useful even though they may have their own limitations. For the next District League Table report in particular, Center for Democratic Development and UNICEF may consider giving a spotlight on M&E besides ranking MMDAs on their development performance. This should allow for in-depth understanding of the M&E challenge and its imperatives for development outcomes. In addition, the M&E Ministry can partner with, for example, the Ghana M&E Forum to develop a M&E specific tool that can serve as a diagnostic tool itself by tracking how MMDAs are executing their M&E functions while identifying M&E champions amongst MMDAs. Here, independent development evaluation units of development partners including multilateral development banks can leverage their vast technical capacity to collaborate with national
evaluation associations and MMDAs in developing such tools.

- **Incentive and recognition:** Recognizing the best MMDAs through some form of an awards scheme is one pathway to encouraging a culture of evaluation. For example, the proposal to develop a tool can serve as a basis for incentivizing and encouraging MMDAs to build their evaluations culture. Ghana M&E Forum’s evaluation awards, which normally recognize institutions and individuals contributing to M&E in the country, is a good starting point. Perhaps, having a dedicated scheme which specifically targets MMDAs should bring more visibility to how they mainstream M&E in their operations.

Besides these recommendations, this review also offers some lessons when it comes promoting a culture of evaluation amongst local governments.

- **Promoting an evaluation culture amongst local governments goes beyond just designing M&E plans.** The enabling environment for actualizing these plans – including financial and technical capacities as well as the right institutional frameworks – should be created.

- **Periodic performance assessment tools such as the District Assemblies Performance Assessment Tool and District League Table in the case of Ghana can be useful in promoting an evaluation culture at the local government level on two fronts.** First, they can help highlight key governance lapses and needs including M&E for which local governments will have to leverage to improve their delivery. Second, such tools have the potential to create greater demand for accountability from communities and civil society.

- **Curtailing the financial challenge that comes with mainstreaming evaluation at the local government calls for the development of evaluation tools that rely intensively on participatory approaches with the contributions of local communities.**

- **The roles of national evaluation associations, research institutions and civil society organizations in general are important for instilling a culture of evaluation.**

- **Mainstreaming evaluation at the local level could be a win-win for improving outcomes for national development priorities and the global goals – the sustainable development goals.** This is crucial at a time when multilateral development banks are looking to engage more sub-nationally.

**Conclusion**

Effective local governance is an essential precondition for quality public service delivery and better development outcomes. A culture of evaluation is equally essential for local government authorities to execute their development planning and administrative functions effectively, efficiently and sustainably while promoting accountability and learning. However, a mixed-bag of factors, including financial, technical and institutional capacities as well as adequate monitoring, are crucial to promoting an evaluation culture amongst Ghana’s MMDAs. It is hoped that the upcoming M&E policy document that specifically target MMDAs will adequately address some of these concerns regarding evaluations.
References


Clement Mensah is a development practitioner and researcher and currently supports energy sector and transport impact evaluations as a consultant at the African Development Bank’s Independent Development Evaluation. His research interests include rural development, sustainable livelihoods, social protection, food security, and territorial development. Prior to his postgraduate studies, Clement worked as a Development Analyst with Development Solutions Centre and served as Monitoring and Evaluation Officer at Genuine Products Ltd., a development consultancy based in Ghana (part-time). He holds Master’s degrees in Development Management and Development Studies respectively from Ruhr University of Bochum, Germany and University of the Western Cape, South Africa. He is currently a PhD candidate in Development Studies at the Institute for Social Development, University of the Western Cape, South Africa.
Interview: Reflecting on the evaluation culture at the African Development Bank
For this edition of Evaluation Matters, we interviewed Armand Nzeyimana, Division Manager, Department for Delivery, Performance Management and Results and Roland Michelitsch, Evaluator General of the AfDB to discuss where the Bank is heading and the prospects for the development of an evaluation culture in the Bank and in Africa in general.

The concept of an evaluation culture has gained some momentum over the last few years, especially with the drive to promote the evaluation of the SDGs. Could you elaborate on what it means for the AfDB?

**Armand:** As we enter the last decade of the SDGs, the stakes could not be higher for Africa. The progress made is certainly encouraging, but clearly insufficient to put the continent on track to meet most of the targets set for 2030. To make matters even more challenging, climate change, persistent conflict and growing inequalities are undermining progress and threatening to reverse gains made in recent years. For the AfDB, the premier development finance institution of the continent, an evaluation culture is critical to understand how our institutional and collective contributions to the SDGs can be enhanced, and more importantly accelerated. Moving away from business as usual to introduce, test, evaluate and scale-up innovative solutions to poverty reduction requires a strong commitment to organizational learning. It also involves greater tolerance for risk-taking and mistakes as platforms for experimenting and learning. For an institution such as the AfDB, striking the right balance between accountability and learning can be challenging. Both values are essential and complementary outcomes of an evaluation culture.

**Roland:** For me, an evaluation culture is one that routinely seeks evidence to continuously improve results. Management guru Peter Drucker coined the phrase “What gets measured, gets done.” So what is really important is to have the right key performance indicators (KPIs) that are systematically measured in a good monitoring system, which for me is also part of an evaluation culture. But KPIs can only take you so far. They give you a first indication of what is improving or deteriorating – but an evaluation culture would then take it further and ask what works, what doesn’t – and in particular why, seeking to improve performance. In the context of AfDB and the SDGs, we should ask ourselves how our operations contribute to reaching the SDGs. For some of them, the answer can come from a relatively simple monitoring system – like how many additional people are getting access to power through AfDB projects, or how AfDB projects change a country’s energy mix to become more sustainable. For others, like ending poverty, it is more complicated. We first need to develop a results framework on how we think AfDB activities are contributing to poverty reduction, both directly and indirectly. Such a results chain will be quite complex, and only good evaluation will be able to test the various hypotheses and assumptions behind that results framework. Ultimately, an evaluation culture will constantly seek to better understand how our interventions contribute to meeting the SDGs, and adjusting our policies, strategies, programs and projects to make an even bigger contribution.
What role do you see evaluation playing in contributing to a stronger results culture in the Bank?

Armand: Evaluation is the healthy pause we take from our daily work to take a hard look at what we are doing right, what we are doing wrong, and more importantly why. It examines key assumptions we make when designing programs, processes, policies and strategies to achieve results, challenging the theory of change underpinning our interventions. In recent years, the evaluation function has examined the progress the Bank is making in implementing the new Development and Business Delivery Model, an ambitious and complex reform agenda seeking to revamp our operational model and organizational structure. The evaluation findings provide critical insights on positive change achieved to date, but also provides essential lessons on how to reap the full benefits of this reform program.

Roland: Evaluation contributes in many ways: It helps to close knowledge gaps, where an institution just doesn’t fully understand the effects of its projects and programs, or the factors contributing to or hampering success. It helps to refine strategies, so that they are better targeted to addressing development needs. Evaluations can also address processes, like IDEV’s evaluations of quality at entry and supervision did, so that we can improve work quality and implement better. There are also monitoring and self-evaluation systems, like those for projects (PCRs or XSRs), that are expected to contribute to a results culture – but only if they are considered reliable. Independent evaluation can check such systems, helping to ensure that they contribute to both accountability and continuous learning. Soon we will evaluate AfDB’s results measurement framework, to see whether it is conducive to continuously improving results.

What relationship between IDEV, governing bodies, and Management of the Bank, do you see as being conducive to strengthening a culture of evaluation and results?

Roland: Ideally, all three actors should always be asking how results have been taken into account – in policies, strategies, processes and operations – and even in the AfDB’s organizational structure. Each of the three parties has a different role, which is elaborated in the AfDB’s Independent Evaluation Policy: (i) the Board provides strategic guidance to IDEV and Management, and ensures that IDEV’s recommendations are acted upon; (ii) IDEV recommends a Work Program that is relevant for the AfDB’s needs, carries out high-quality evaluations and disseminates the findings and lessons in a format that is appropriate for the respective audience (be it the Board, Management, or project staff) – and thus helps the AfDB improve; and (iii) Management takes into account the lessons, develops appropriate action plans and adjusts policies and operations to improve results. Between IDEV and Management, it is important to have a collaborative approach – after all, we all want to improve development results for the people in Africa – while at the same time respecting each other’s role: IDEV’s role in providing an independent assessment and broad guidance through its recommendations (e.g., WHAT should be replicated or needs to be improved), and Management’s role in managing and operationalizing the improvements through its actions (e.g., HOW to make it happen).
What about evaluation culture in African countries? Where do you see opportunities to promote this evaluation culture in Africa from 2020? Where are the challenges?

Armand: In the context of ambitious social and economic aspirations aligned to the SDGs and the Africa 2063 Agenda, more and more countries on the continent recognize the need to apply a focused and disciplined approach in evaluating the effectiveness of national transformation plans. The drive to accelerate progress on the SDGs during this next decade provides a unique opportunity to support this effort and promote an evaluation culture across the continent. In particular, there is a growing appetite among decision makers for tracking and evaluating public sector performance as evidenced by the recent interest in setting up Delivery Units at the Center of Government – Presidency or Prime Minister’s Office. Indeed, Delivery Units respond to rising public expectations for faster results, tangible benefits, effective accountability in the use of resources and greater value for money. This is an area where the Bank is taking an increasingly active role in supporting these efforts. The most significant challenge in promoting an evaluation culture on the continent remains poor data quality and weak data management systems, an issue requiring our urgent attention.

Roland: Indeed, African leaders increasingly want to promote an evaluation culture. In part, this is driven by more information flows (citizens have greater information about their own and their country’s situation as compared to others), and growing demands by the people they govern to tangibly improve their lives. So evidence of positive results is becoming more valuable to help them or their parties to stay in power. IDEV has contributed to promoting an evaluation culture in African countries by both stimulating demand for evaluations – hosting the secretariat of the African Parliamentarians’ Network on Development Evaluation (APNODE) for example – and fostering a greater supply of evaluations – by supporting Twende Mbele, a south-south government peer learning initiative aimed at building monitoring & evaluation capacity and sharing experiences. In my experience, it is relatively easier to promote interest in evaluating and demonstrating what works well. Admitting things that don’t work well is usually a much greater challenge – both in the AfDB and on the continent. But learning from failures is also essential for an evaluation culture that seeks continuous improvement.
A new generation of advocates: young and emerging evaluators and the promotion of a global youth-participatory evaluation culture

Since the International Year of Evaluation in 2015 (EvalYear) and the launch of the Global Evaluation Agenda (EvalAgenda 2020) in 2016, Young and Emerging Evaluators (YEEs) have progressively conquered a place within the global evaluation community and have increasingly contributed to shaping the global evaluation culture. In doing so, they have positioned themselves as a new generation of advocates for what we could call a “youth-participatory evaluation culture”, one that harnesses the values and skills that youth and YEEs can bring to the evaluation table, and where the role of young evaluators is defined as one of collaborator and co-leader. We believe that such a culture should be promoted in 2020 and beyond, and we call on all African YEEs who identify as professionals and who value social change to join forces with us. Together, we can fuel a movement towards the achievement of a global youth-participatory evaluation culture.
Key Messages

- YEEs are gaining an increased interest within the global evaluation community; multiple national, regional, and international Voluntary Organizations for Professional Evaluation (VOPEs) are pursuing initiatives to bring the ideas and perspectives of evaluation novices into their work and into the global evaluation culture.

- YEEs have organized themselves to take ownership of the EvalAgenda 2020 and have brought specific youth values and skills to the table; they are now pushing for increased youth participation in evaluation and advocating for a move from evaluation on youth to evaluation with and by youth.

- In the era of the Sustainable Development Goals (SDGs) and youth activism, the emergence of a youth-participatory evaluation culture holds interesting opportunities for African YEEs, who could consider exploring ways to build a ‘Made in Africa Evaluation’ that incorporates principles of ‘youth-participatory evaluation’.

Introduction

In 2014, EvalPartners, the global movement to strengthen national and civil society’s evaluation capacities, started a global, multi-stakeholder consultative process to brainstorm on the development of a Global Evaluation Agenda for 2016-2020, which we now call “EvalAgenda 2020” (EvalPartners, 2016). This document, which is believed to be the “first ever long-term global vision for evaluation”, defines priorities for evaluation during the first five years of the 15-year period addressed by the new SDGs. The United Nations calls for a rigorous review of progress towards the global goals, “based on evidence and informed by country-led evaluations” (United Nations, 2015), and the EvalAgenda 2020 promotes the establishment of this strong enabling environment for evaluation, in order to achieve a “positive, broad-based global evaluation culture”. In doing so, it invites all stakeholders to take “widespread ownership of the agenda”, building upon the idea that evaluators should become global advocates for their profession.

In this article, we suggest that African YEEs have organized themselves to take ownership of the EvalAgenda 2020 and play an increasing role in the promotion of a global evaluation culture on the continent. In doing so, YEEs have positioned themselves as advocates for the professionalization and the institutionalization of evaluation as an essential management and social change instrument in the SDGs era. We assert that African YEEs, along with peers from across the globe, have contributed to shaping the current global evaluation culture, by bringing specific values and skills to the table and by pushing for increased youth participation in evaluation (Montrosse-Moorhead et al., 2019). As a result, we may be witnessing the emergence of what can be called a “youth-participatory evaluation culture” and we invite African YEEs to remain engaged in order to further contribute to the shaping and strengthening of such an evaluation culture in 2020 and beyond.
“Engage, Innovate, Exchange”: integrating the “Youth and Evaluation” nexus into the global evaluation culture

Launched by EvalPartners in 2015, EvalYouth is a global network established to promote engagement, innovation and exchange on evaluation among YEEs, youth stakeholders and other key actors. It was created with the aim to address two major ongoing issues facing the global evaluation community: one, the need for appropriate representation and participation from all stakeholders, including from youth (who are frequently absent from evaluations of policies and programs that impact them); and two, the need to build a pool of skilled evaluators in order to secure the production of high-quality evaluations of SDG-related initiatives. EvalYouth’s overall purpose is to “support YEEs to become technically sound, experienced and well-networked professionals who contribute to evaluation capacity at national, regional and international levels” (EvalPartners, 2016). The network has therefore contributed to highlighting the presence of YEEs in the evaluation field, focusing on work to ensure that capacity development efforts are aligned to the needs and challenges faced by newcomers to the evaluation community.

YEEs are either young evaluators under the age of 35, and/or new evaluators with less than 5 years of experience in evaluation, and/or recent university graduates who are willing to join the evaluation profession, and/or development professionals who have technical knowledge on evaluation and are willing to become evaluation professionals. YEEs enter a robust field, one that requires the acquisition of the specific knowledge, skills and dispositions needed to achieve a standard that constitutes high quality evaluations (Stevahn et al., 2005; Tourmen, 2016). Their professional growth and development entail a continuous learning process over several years of practice, from novice, to emerging to a proficient level.

The EvalAgenda 2020 defines three pillars for the cultivation of evaluation capacity by EvalYouth. All play a role in the achievement of a positive, broad-based global evaluation culture:

- the ENGAGE pillar (political dimension) targets impact through social mobilization with the objective to enable VOPEs,


governments, policy makers, civil society representatives and advocates, international development agencies, academic institutions and other interested actors to coordinate efforts to better engage YEEs in the evaluation field and young people in the evaluation process;

- the INNOVATE pillar (entrepreneurial dimension) targets impact through innovation in practice with the objective to facilitate innovation in the evaluation process by engaging the community of policy makers and evaluators to make use of new approaches, strategies and methodologies that can attract and take advantage of the ideas and energies of YEEs and young people;

- the EXCHANGE pillar (social dimension) targets impact through learning and knowledge sharing with the objective to serve as a platform to share knowledge, learning and experiences on the best ways to strengthen the engagement of YEEs in the evaluation field and strengthen the inclusion of young people in the evaluation process.

The launch of EvalYouth has sparked an unprecedented interest in YEEs within the global evaluation community. In the past five years, multiple national, regional and international VOPEs have launched initiatives to bring the ideas and perspectives of YEEs into their work. On the African continent, several YEE networks have gradually been established, of which two regional YEE networks have embraced the EvalAgenda 2020 and embarked on assisting their VOPEs in the promotion and achievement of a global evaluation culture. These two networks are affiliated to the African Evaluation Association (AfrEA) and the Francophone Network of Evaluation, or Réseau Francophone de l’Évaluation in French (RFE). They are both considered as regional chapters of the wider EvalYouth global network and operate with action plans that are aligned to the three pillars described above.

Along with other peer networks from different continents, these two African YEE networks have progressively organized themselves to constitute a global advocacy coalition for the promotion of shared beliefs on the crucial need to gravitate towards a renewed vision of what a global evaluation culture should look like. In doing so, they have contributed to the rise of what four of the founding members of the EvalYouth
A new generation of advocates: young and emerging evaluators and the promotion of a global youth-participatory evaluation culture

In their enlightening and somewhat provocative piece recently published in an IDEAS flagship publication on evaluation and transformational change, our EvalYouth colleagues define youth-participatory evaluation as: “the process of involving young people in conducting evaluations. It includes the participation of young people in different roles (evaluation managers, evaluators, evaluation reference group members, and informants) in all phases of an evaluation (preparatory phase, design phase, data collection phase, analysis and reporting phase, and facilitation of use and dissemination phase)“. The purpose of such an evaluation approach is: “to empower young people, to recognize their potential, and to acknowledge their legitimate and unique perspectives by meaningfully and sustainably engaging them in evaluation and by focusing on issues that affect their lives. In short, it moves evaluation from being framed as something that is done to young people to evaluation with or by young people” (Montrosse-Moorhead et al., 2019).

The implications of such an approach in terms of evaluation practice are summarized in the following illustration.

What does this mean for African YEEs willing to further engage in advocating for the promotion of a global evaluation culture that would value the full potential of young people and would grant YEEs a seat at the table as equal partners?

We believe that YEEs should consider:

- Sharpening their advocacy skills in order to influence decision-makers to increase the demand and supply for youth-participatory evaluation. As stated in a flagship toolkit on advocating for evaluation (Karkara and al. 2014), “establishing an enabling environment for evaluation is as much a political exercise as an issue of developing technical
Figure 4: Modalities for youth participation in evaluation

01 Business as usual
Young people as informant FOR an evaluation
- Limited youth engagement
- Youth are consulted and give feedback
- Youth have no power in the decision-making process

02 Sometimes
Young people as advisors FOR an evaluation
- Specific youth engagement
- Youth work with evaluators and contribute to a specific area
- Youth do not contribute to decision-making, except for that specific area

03 Rarely
Young people as co-evaluators IN an evaluation
- More important youth engagement
- Youth and seasoned evaluators work as a team to envision, design and implement
- Youth share decision-making power throughout evaluation process

04 Almost never
Young people as leaders OF an evaluation
- Maximal youth engagement
- Youth envision, design and implement and seasoned evaluators coach and facilitate
- Youth hold all of the decision-making power throughout evaluation process

Source: Designed by the authors based on the data included in Montrosse-Moorhead et al. (2019) and Checkoway B., Richards-Schuster K. (2003).

As we currently have the largest generation of young people in human history, with African youth aged below 35 now representing nearly 75% of Africa's population, members of YEE networks should build their ability to advocate for the mainstreaming of youth-participatory evaluation. They can do so by drawing from a substantial body of literature that will nurture their narrative regarding the benefits of youth-participation for evaluation quality, for young people as individuals, for the communities in which youth reside, for organizations that conduct or commission youth-participatory evaluations, etc.

Exploring ways to build a ‘Made in Africa Evaluation’ that incorporates principles of youth-participatory evaluation. Montrosse-Moorhead et al. (2019) have underlined that because of their values, YEEs favor an evaluation culture that seeks diversity and inclusivity and ‘are not methodologically dogmatic’ but rather “tend to embrace mixed methods and contextually responsive approaches”. As we are seeing the rise of culturally responsive evaluation and indigenous perspectives in evaluation, we believe that YEE networks in Africa should consider joining the call for a ‘Made in Africa Evaluation’ by offering a viewpoint from the youth perspective on the local values, assumptions and practices that should underpin this concept.

Accelerating exchange and transfer of innovative ideas between YEE networks across the African continent. We believe that African YEEs have all the necessary components in hand.
to achieve that acceleration, as they possess skills that are specific to youth, such as technological savviness and connectivity, and can easily use social and media outlets to collaborate in a more sustained way. We encourage YEE networks to build stronger alliances amongst themselves as well as with other equity and social justice-focused groups operating at their local, national or regional levels. We hope they can find cutting-edge technological solutions to the language barriers they might face in order to exchange from one sub-region of the continent to the other, and we also hope to welcome more lusophone YEEs within the African YEE movement (De Morais Sarmento et al., 2019).

Conclusion

As much as young people constitute a specific segment of the global society, YEEs are now starting to form a specific segment of the global evaluation community. As we move towards 2020 and beyond, YEEs will bear an increasingly important responsibility for the evaluation of interventions related to the SDGs. The United Nations public outreach website dedicated to the SDGs lists five potential roles young people can play in the overall effort to achieve these global goals. We would like to call on all African YEEs to consider joining the movement to further fuel the “revolution in the making”, and, by doing so, to reflect on the type of role they might want to play as young evaluators for the development of advocacy strategies that will contribute to strengthening an enabling environment for evaluation based on a youth-participatory evaluation culture. Will they be critical thinkers, who identify and challenge existing power structures and barriers to change, and expose contradictions and biases? Will they be change-makers, engaged in youth-activism and the mobilization of their peers and other stakeholders? Will they be innovators, offering new ideas and alternative solutions? Will they be communicators, spreading ideas to their peers and communities at the local level, as well as across countries and regions? Will they be leaders, driving change in their communities and countries? Or will they even go beyond the scope of this frame?

We are highly confident that the growing pan-African YEE movement, stirred by the power of group genius and collaboration, will be a major contributor to the achievement of a global youth-participatory evaluation culture in 2020 and beyond.
Endnotes

1. This article draws on two presentations delivered during a panel session on YEEs held at the International Development Evaluation Association (IDEAS) Global Assembly in Prague in October 2019. During this session, representatives of five YEE networks affiliated to different regional VOPEs reflected on the extent to which their activities have contributed to the realization of the EvalAgenda 2020 objectives, and on the opportunities that exist or can be created for YEEs to increase youth participation in the field of evaluation. We delivered our presentations respectively on behalf of the Francophone Network of YEEs (Réseau Francophone des Évaluateurs Émergents) and of the AfrEA Network of YEEs. We were both granted an IDEV/AIDB bursary to attend this Global Assembly, respectively as Northern and Southern African YEEs. We wish to express our sincere gratitude and thanks to IDEV and IDEAS for their support.


3. AfrEA is a regional umbrella organization for African national VOPEs. It serves as a leading source of knowledge for individual evaluators in countries where national evaluation associations do not exist. www.aorea.org

4. RFE is a transnational umbrella organization for francophone national VOPEs located in both the Global North and South (North America, Western Europe, North Africa, West Africa). Its purpose is to strengthen evaluation capacities in the Francophone area and develop a technical evaluation knowledge corpus in French. www.rferevaluation.org

5. To name a few regional networks: EvalYouth Latin America and Caribbean, EvalYouth Eastern Europe and Central Asia, EvalYouth Asia, YEE network of the European Evaluation Society. These regional networks are supplemented by several national YEE networks affiliated to national VOPEs all over the world. At the global level, IDEAS has also recently integrated a YEEs Thematic Interest Group within the scope of its work.

6. Designed by authors based on their experiences as active members of the management groups of these two YEE networks.


8. Mentioned by Montrosse-Moorhead et al. (2019). A detailed description of these roles can be found here : https://www.un.org/sustainabledevelopment/youth/

References


Fazeela Hoosen is a director and evaluator at Benita Williams Evaluation. Ms. Hoosen has experience in undertaking utilization-focused evaluations in the education sector, which involve qualitative and quantitative analysis. She recently directed her focus to providing Monitoring and Evaluation technical assistance to and evaluating governance networks that work across Africa. Ms. Hoosen holds a Master’s Degree with a Post-graduate Diploma in Monitoring & Evaluation. She served a three-year term (2014 – 2017) as Board member of the South African Monitoring and Evaluation Association (SAMEA). Ms. Hoosen is involved in the global EvalYouth movement as well as the regional AfrEA YEE Network to support and create opportunities for YEEs to build and strengthen their evaluation practice.

Kenza Bennani is an evaluation manager with the French Ministry of Foreign Affairs, specializing in international development and cultural diplomacy policies. She is actively engaged in international (IDEAS, RFE), regional (AfrEA, EES) and national VOPEs (SFE, AME). She is currently a board member of the International Development Evaluation Association (IDEAS) representing the Middle East and North Africa, a board member of the Francophone Network of Emerging Evaluators (RF-Eé), a stream leader of the IDEAS Thematic Interest Group for Young and Emerging Evaluators (YEE ITIG) and the co-founder and co-convener of the French Evaluation Society (SFE) working group for Young and Emerging Evaluators (Jeunes Evaluateur.rices Emergent.es). Prior to being an evaluator, Ms. Bennani worked as a public sector management consultant for Capgemini Consulting, and also gained experience in the areas of advocacy, public outreach and private-sector fundraising with several UN system and Red Cross Movement organizations. With a multi-faceted academic background, Ms. Bennani holds an M.Phil in Public Policy and Administration from Université Panthéon-Assas, an M.A. in International Sustainable Development Geography from Université Paris-Sorbonne and a Master in Management from Audencia Business School.
Faced with internal and external pressures, member countries of the Organization for Economic Co-operation and Development (OECD) have developed a culture of evaluation tied to the implementation of public policies and programs. Unlike Western countries, African countries are still building their evaluation culture. However, several initiatives have been undertaken in different African countries including Benin, Côte d’Ivoire, Senegal, Uganda and South Africa through the establishment of national evaluation systems or mechanisms and tools favoring the use of evaluation in the implementation of public policies. Such a culture is a key part of good governance and is fostered by the existence of a number of opportunities. Nonetheless, these countries face certain threats in their quest for creating an evaluative culture. These include a weak capacity in statistics, program monitoring and the challenge of evaluating certain public services. This article discusses what a culture of evaluation involves and what can be done to create and maintain such a culture. It focuses mainly on African countries. It also aims to promote the culture of evaluation in Africa.
Dékaï Jean-Luc Namou, Economist, Ministry of Health and Public Hygiene of Côte d’Ivoire and Marie Laure Essis, Head of Monitoring and Evaluation Service National Institute of Public Health of Côte d’Ivoire

Key Messages

- The development of a culture of evaluation requires deliberate efforts by a country and strong leadership to encourage and support such a culture.
- Developing countries face challenges similar to those faced by developed countries in moving towards a culture of evaluation.
- For a culture of evaluation to be established and implemented in any country, interested stakeholders and a commitment to transparency and good governance are necessary.

A culture of evaluation: different approaches in culturally diverse regions including Africa

To develop effective evaluation functions, countries need a strong evaluation culture in which evidence is deliberately sought in order to allow better implementation and execution of their development programs. Without such a culture, any effort to put in place effective evaluation activities would be compromised (Stewart, 2014).

Developing a culture of evaluation not only involves good intentions and osmosis (Mayne, 2010), but also requires deliberate efforts by a country as well as strong leadership to encourage and support such a culture. To do this, managers and staff need to know that there is clear political will at the country level and that information based on results and evidence is valued. This should be an integral part of planning, budgeting, execution and review chains (ibid.).

A mapping of the evaluation cultures of OECD countries was carried out in 2002 by examining 23 countries on nine variables relating to: evaluation in policy areas, the availability of specialized evaluators, the existence a national discourse on evaluation and an evaluation profession, and the establishment of institutional mechanisms for conducting evaluations and sharing their results. The mapping found that Australia, Canada, the Netherlands, Sweden, and the United States have the highest "rating of evaluation culture" among OECD countries (Furubo, Rist and Sandahl, 2002, cited in Imas and Rist, 2009).

OECD countries have developed a culture of evaluation in response to varying degrees of internal and external pressures. France, Germany, the Netherlands, Australia, Canada, the Republic of South Korea and the United States, for example, have developed a culture of evaluation largely due to strong internal pressures such as national planning, programming and budgetary requirements for new socio-economic spending programming, as well as legislative monitoring (Imas and Rist, 2009). Many of these countries were better equipped to do so because they have democratic political systems, strong empirical traditions, civil servants trained in the social sciences (as opposed to legal training), and efficient institutional and administrative systems (ibid.). Countries with high spending on education, health and social protection have adopted evaluation mechanisms ➤
in these areas, which then influence other areas of public policy. These pioneer countries have contributed to the dissemination of a culture of evaluation in other countries through evaluation organizations, training institutes, networks and consultancy firms (Imas and Rist, 2009). By contrast, some countries such as Ireland, Italy and Spain have adopted a culture of evaluation due to strong external pressures, mainly from the conditions of membership of the European Union (EU) including access to EU structural development funds. The latter were strongly influenced by the culture of evaluation of pioneer countries, as well as by the culture of evaluation rooted in the international community with which they interact (Imas and Rist, 2009).

In 2009, most of the thirty OECD countries had mature evaluation systems. However, establishing them was not an easy nor linear process. Three approaches enabled them to set up their systems. These are: an approach that focuses on the whole of government, one that focuses on parts of government and a mixed approach. The first involves extensive and comprehensive implementation of the system across government and requires time, at least a decade (in the case of Australia). The second approach focuses on a part or sector of government, such as a single ministry (for example, Mexico focused on social development, Jordan on planning and the Kyrgyz Republic on health). The third approach integrates the first two (the case of Ireland) (Imas and Rist, 2009).

**Development of evaluation in Sub-Saharan Africa**

Unlike OECD countries that already have sophisticated governance systems, those in Africa are just beginning to build an evaluative culture. The majority of external evaluations carried out in most African countries have been due to the requirement of donors (Tinsakré, 2013). According to Mouton and Wildschut (2017), donors remain the primary source for the request of evaluations on the African continent. However, this situation is not likely to encourage the effectiveness of public policies and the judicious use of state resources. Furthermore, the absence of a strong and internal political will to develop a culture of evaluation prevents its development.

Some countries have developed national evaluation policies and plans, of which the example of South Africa is very pertinent. Citing the case of Benin, Uganda and South Africa, Goldman said, “given their limited resources and capacities, the three countries started their respective National Evaluation Systems with national evaluation priorities indicated in an evaluation plan or program, using donor resources if necessary, but piloting the program themselves. This is important if the evaluation should be part of the countries’ strategic agendas, and not just recommended by the donors” (Goldman et al., 2018, cited in Evaluation Matters 2018: 29-30).

Benin, Côte d’Ivoire and Senegal have adopted mechanisms and tools that promote the use of development evaluations in the implementation of public policies (IPU, 2009, cited in Evaluation Matters, 2018: 31). In addition, since 2010, Benin, Uganda and South Africa have made considerable efforts to integrate evaluations into the work of public authorities, starting from very different political situations and with different resource constraints. There has been great local innovation in the establishment of these systems as well as evolutionary management as these systems develop. According to their different economic policies, South Africa undertook 56 evaluations in December 2016, compared to 23 for Uganda and 15 for Benin, reflecting a greater capacity of governments to fund evaluations. However, the scope of these evaluations differ from country to country.
Benin’s evaluations in particular are at the policy level rather than the program level, thus covering a broader field, but in a less in-depth manner (Goldman et al., 2018, cited in Evaluation Matters, 2018: 24-25).

The case of South Africa presents a good working example where a national evaluation system was implemented in 2012 and is headed by the Planning, Monitoring and Evaluation Department (DPME), previously the Ministry of Performance Monitoring and Evaluation. However, even while the framework for the national evaluation policy was being developed, the country pursued some pilot evaluations. The initial concept of an evaluation of results in addition to a devoted research unit (see Box 1) to drive the evaluation system was approved in September 2011. The aim was “to coordinate the evaluation function within government, ensuring that the high-quality evaluation and research underpin public policy and programming, to maximize the impact of government and services” (Goldman, 2012, cited in African Evaluation Journal Vol 3, No 1). The first pilot evaluation started in October 2011 and was completed in June 2012. The National Evaluation Plan (PNE) was approved by cabinet in June 2012, with eight evaluations performed (DPME 2012a). The 2013-2014 PNE was approved in November 2012 with fifteen evaluations (DPME 2012b) and the 2014-2015 PNE, approved in November 2013, saw fifteen evaluations (DPME 2013b) (Goldman, 2012).

**BOX 1. KEY ROLES FOR A RESEARCH UNIT (GOLDMAN 2012)**

- Development and maintenance of an evaluation policy framework within government, as well as a three-year and annual evaluation plan.
- Construction of a whole-of-government approach to advance evaluation and ensure the use of evaluations to inform plans and budgets.
- Development of technical specifications, systems and guidelines for evaluation within government.
- Implementation and support of evaluations and research.
- Monitoring and quality control of a government-wide evaluation process.

Importance of a culture of evaluation

Other terms used for a culture of evaluation include a culture of results and a culture of performance (Mayne, 2010). In either case, such a culture uses results for decision-making and learning and is an adequate means for the implementation of a results-oriented policy (DAC, 2014). In addition, it promotes the use of information on results to improve organizational performance and is thus used for planning, monitoring progress, learning and decision-making (DAC, 2014). Information on results can also be used for reporting purposes or as a management tool. In the area of performance management, these are used to improve organizational performance and achieve better results. (DAC, 2014). Members of the Development Assistance Committee (DAC) use them primarily for external reporting for accountability and communication purposes, monitoring progress, and decision-making. Organization wide, performance information is primarily used for communication with the outside world. At the strategic management level, information on results is used not only for accountability purposes, but also to inform decision-making and lessons learned. At the level of projects or programs, they are used firstly to monitor progress and secondly to make decisions (DAC case, 2014). The culture of evaluation is key to building more effective approaches to managing and evaluating results (Mayne, 2010).
Why develop a culture of evaluation in African countries?

Results-based management must aim to foster a culture of performance, hence a culture of evaluation (Mayne, 2010). Monitoring and Evaluation are a key part of this process. Monitoring is the ongoing part of project/program management, which involves the systematic collection and analysis of data based on intervention indicators. It is an internal assessment aimed at keeping interventions on track and ensuring the timely decision-making necessary to improve their design and operation. Outcome evaluations are essential for shedding light on gaps, determining when adjustments need to be made and what types of changes are required - change of focus and/or adjustment of the project/program strategy (theory of change, logical framework and/or implementation strategy), redefinition of priorities, capacity building, allocation of resources more effectively, etc. (Manual for Results-Based Management and the 2030 Agenda for Sustainable Development). Thus, according to the Development Assistance Committee (DAC, 2014) the potential consequences of the absence of an evaluation culture can be summarized as follows:

- Difficulty in highlighting positive changes in expenditure allocated to development aid.
- Insufficient evidence to inform decision-making and lessons learned.
- Difficulty demonstrating the results of an intervention.
- Low valuation of performance information.
- Obstacle to the use of results for decision-making and learning.
- Lack of staff motivation, leadership and incentives.

In light of the above, the lack of an evaluative culture undermines attempts at good governance, transparency and accountability. It is therefore of paramount importance for African countries to support a culture of evaluation through the strengthening of good governance, democracy, improving the quality of public services and accountability.

In the quest for a culture of evaluation in Africa, there are several threats, possibilities and opportunities, which are summarized in Figure 1 [CAD, 2014; Dejean J. et al., 1998; Nasser, 2015; UNODC, 2019]

To progress towards a culture of evaluation, African countries must confront these threats and provide adequate solutions. The above favorable circumstances are an asset to achieve this.
But, how does one recognize an evaluative culture? What are its key features?

**Characteristics of a functional culture of evaluation**

A culture of evaluation refers to a culture where empirical performance information is deliberately researched in order to use it to learn how to better manage programs and services, and thereby improve performance. Such a culture values empirical evidence on results. A number of authors and reports have examined the issue of an evaluation culture and how to achieve it. A synthetic presentation of the different results obtained by the studies carried out on the culture of evaluation highlights its characteristics in three main points (Mayne, 2010).

Firstly, a culture of evaluation deliberately looks for evidence of achievement through monitoring and evaluation. It is therefore evident in a self-assessment. Secondly, it engages in evidence-based learning, in other words, it promotes the creation of learning milestones, learning from mistakes made and poor performance; and encourages knowledge transfer. Thirdly, it encourages experimentation and change, that is to say, it supports risk-taking and searches for new ways of working (ibid.).

**Recommendations for developing a real culture of evaluation**

To encourage the development of an evaluation culture, it is recommended that African governments set up the minimum preconditions presented in Figure 2 below:

**Conclusion**

Evaluation allows countries to have confidence in their development programs, helping them evolve successfully and use evidence to guide their decisions. The promotion of a culture of evaluation...
across organizations in 2020 can only take place through an assessment of organizational readiness, planning for cultural change, capacity building, and strengthening of organizational support systems. Only these efforts can develop and create a flourishing evaluation culture.

The existence of national evaluation societies as well as the African Evaluation Association (AfrEA), whose vision is an Africa rooted in a culture of evaluation for equitable and sustainable development (AfrEA, 2017), are important assets that governments can rely on. AfrEA, which is increasingly working to promote "Made in Africa" evaluations, is ready to work with African governments to increase knowledge and use of evaluation approaches within public institutions. However, strong and sustained government leadership is paramount to achieving this goal.

1. Health Economist, Ministry of Health and Public Hygiene, Training and Research Directorate, Côte d’Ivoire. Email: jeanlucnamou@gmail.com/ namoujeanluc@gmail.com

Endnotes
References


Kone, N.N. and Tioyé, S.A.M. (2018).“Défis et enjeux de la promotion de l’évaluation du développement au niveau législatif” EVALUation Matters, 1st quarter 2018, AfDB.


Jean-Luc Dékaï is a health services engineer specializing in health economics at the Ministry of Health and Public Hygiene (MSHP) in Côte d’Ivoire since May 2019. Before joining the MSHP, he was responsible for coordinating projects at Groupe Conseil Santé et Développement Sud, a consultancy firm based in Abidjan. He is an active member of the Ivorian Initiative for Evaluation (iiEval). He holds a DESS in Health Economics from the African Center for Higher Studies in Management, and a Masters in Food Microbiology from the University of Abobo-Adjamé in Abidjan (newly named the Nangui Abrogoua University).

Dr. Marie Laure Essis is a Public Health Doctor, Infectiologist and Research Associate. Dr. Essis is the Head of Monitoring and Evaluation Service at the National Institute of Public Health in Abidjan, Côte d’Ivoire (INSP). She also holds a Masters of Research – specializing in Planning and Management of Health Programs and Services from the African Institute of Public Health (IASP) in Ouagadougou (Burkina Faso) and a Diploma of Advanced Study in Public Health – specializing in Health economics from the UFR of pharmaceutical sciences in Abidjan (Côte d’Ivoire). Dr. Essis is currently working towards a PhD in Public Health at IASP. She also coordinates and facilitates training in Monitoring and Evaluation organized by INSP. She has extensive experience in project/program evaluation.
This article discusses the dramatic changes that will face evaluation practice and the “culture of evaluation” in Africa as the 4th Industrial Revolution [4IR] evolves. The drivers of the 4IR will be the rapid evolution of six “disruptive technologies”, five of which present great opportunities for evaluation practice, but also significant challenges. The culture of evaluation will be transformed, and the key questions are how well the transformation will be managed and whether evaluation will remain a major force in the new big data ecosystem.
Michael Bamberger, Independent Evaluation Consultant

Key Messages

- The Fourth Industrial Revolution is developing powerful new information technologies that can significantly strengthen how evaluators collect and analyze data.

- However, these technologies will inevitably be disruptive and it will be a challenge for the “culture of evaluation” to adapt and to incorporate these new technologies.

The culture of evaluation

It is helpful to begin by defining what we mean by the “culture of evaluation”. Mayne (2008) argues that an evaluative culture denotes an organizational culture that deliberately seeks out information on its performance in order to use that information to learn how to better manage and deliver its programs and services, and thereby improve its performance. An organization with a strong evaluative culture engages in self-reflection and self-examination as well as evidence-based learning, and encourages experimentation and change. Mayne also discusses ways to develop an evaluative culture, focusing on leadership, organizational support structures and a learning focus.

While this provides a helpful framework, it is important to recognize that evaluations are conducted for many different purposes, by different kinds of organizations and at different levels – ranging from small projects in a few villages, to a 15-year global evaluation program for assessing the effectiveness of the Sustainable Development Goals (SDGs). Evaluations are also conducted by various organizations operating with different mandates.

Evaluations are likewise conducted for varied purposes, with agencies giving different priorities to differing purposes. A widely used classification identifies three main types of evaluation: developmental evaluation (designing and testing innovative approaches for addressing problems, particularly for understanding programs operating in complex contexts); formative evaluation (helping managers and planners to improve the design and implementation of an ongoing intervention or to learn lessons that can improve future interventions); and summative evaluation (assessing the overall merit, worth and significance of a program). However, evaluations may have a range of other purposes.

Another important dimension concerns the evaluation design. At least ten main types of evaluation design have been identified, each involving a unique methodology with differing sets of assumptions about broader considerations such as the nature of evidence and whether there are objective “facts” that remain constant for different observers (epistemology and ontology). These designs are listed in Box 1. The list is divided into current evaluation designs and new, evolving designs adapted to the new reality of the Fourth Industrial Revolution (4IR). However, there is a continuum and some current evaluation designs are starting to address these new realities.

So, given the fundamental differences between many of these methodologies, as well as the diverse purposes of...
The culture of evaluation in the age of big data: The need for a new evaluation paradigm for the 4th Industrial Revolution?

Many evaluation approaches focus on continuity and are resistant to change

An important element of what could be called “mainstream evaluation” (the approaches used by many United Nations, multilateral and bilateral development agencies and which are taught in many evaluation training programs) are the demands for standardized evaluation approaches which are widely applied and which only change slowly over time. There are a number of factors influencing continuity.

Methodological considerations: Mainstream evaluation theory argues that the best way to assess causality and impact is to obtain a baseline measurement on outcome indicators and then to repeat the measurement using the same sample and indicators at the completion of the project. Ideally, the same data should be collected for a matched comparison group. This methodology requires that identical indicators be used at different points in time. Agencies that are required to evaluate all their programs will often invest significant resources in developing and testing the measurement instruments and in 

<table>
<thead>
<tr>
<th>BOX 1. CATEGORIES OF EVALUATION DESIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current evaluation designs</strong> (all of which can be adapted to the 4IR)</td>
</tr>
<tr>
<td>1. Experimental designs</td>
</tr>
<tr>
<td>2. Quasi-experimental designs</td>
</tr>
<tr>
<td>3. Statistical designs (econometric)</td>
</tr>
<tr>
<td>4. Theory-based evaluations</td>
</tr>
<tr>
<td>5. Case-based evaluation – including Qualitative Comparative Analysis</td>
</tr>
<tr>
<td>6. Participatory and qualitative evaluation</td>
</tr>
<tr>
<td>7. Gender responsive evaluation</td>
</tr>
<tr>
<td>8. Synthesis and review</td>
</tr>
</tbody>
</table>

Big data: geospatial, internet of things, social media, smartphone generated data

Smart data analytics: Dashboards, integrated data platforms, predictive analytics, digital experimentation, machine learning, artificial intelligence

selecting the sample. Consequently, there are incentives to continue to use the same instruments to ensure comparability of results. To a critic this can be seen as a conservative mentality, which is resistant to introducing new instruments and indicators to adjust to change.

Organizational and institutional considerations: Many agencies conduct large numbers of evaluations of similar programs and need a standardized approach to reduce time and cost as well as to ensure quality. For example, agencies that work at the global or regional level are required to conduct periodic evaluations of their country programs, or sometimes to evaluate each of their individual loans/investments. Often evaluations will be sub-contracted to consultants, and consequently standardized methodologies, procurement procedures and terms of reference may be required.

For some national evaluation agencies, transparency is a priority. Particularly where evaluations are used as part of government decision-making to determine resource allocation among different states or communities, it is important to develop consensus on the methodology, and to provide stakeholders with access to the methodology.

As we will discuss below, these methodological and organizational factors contribute to the slowness of many evaluators to adopt potentially powerful new tools and techniques for data collection and analysis that are becoming available through the 4IR.

The culture of evaluation in the 4IR: The need for a new evaluation paradigm

A 2019 AfDB study, “Unlocking the potential of the 4th industrial revolution in Africa,” identified six potentially disruptive technologies that will dramatically change economic development patterns in Africa: Artificial intelligence (AI); the Internet of Things (IoT); Big Data; Blockchain; drones and 3D printing. In parallel to the dramatic transformations that these technologies are already producing in industry and commerce, they are also beginning to influence how development programs are designed, implemented and monitored.

All of these technologies dramatically increase the kinds of data that are available for the design, implementation and monitoring of social and economic development programs, and potential reduce the costs of access to and analysis of these data. This includes the more precise targeting of products and services to particular audiences, and the early detection of problems with project implementation of intended outcomes.

The rapid evolution of these information technologies and the increasing range of their applications has very important implications for development evaluations. There are three main challenges:

- How to assess the impacts of these new information technologies on economic and social development – including who benefits and who is excluded or may even be worse off;
- How to assess the impacts of these technologies on development policies and programs; and
- How to assess where, to what extent and how these new technologies can be incorporated into the design and implementation of development evaluations.

Initial impressions suggest that these kinds of questions concerning evaluation receive very little attention in many discussions on the 4IR. For example, the 2019 AfDB study mentioned above focuses on the business case for the promotion of these new technologies but does not discuss
how to evaluate the effectiveness and impacts of these technologies, nor is there any discussion of unintended consequences. We will argue in the following sections that the implications of these new information technologies are so profound that they will require a paradigm shift in the culture of evaluation.

The challenges facing most current evaluation approaches, and which will be exacerbated by the 4IR

Evaluators are aware of the methodological, economic and organizational challenges that already confront the evaluation of current development programs. Many of these challenges will become more critical for evaluation in the 4IR era, and other new challenges will emerge. Some of the current evaluation challenges, all of which will continue in the 4IR era include:

Cost of data collection: For many evaluations, data collection represents the largest cost. Each interview involves a significant cost, therefore creating pressure to reduce the sample size or to exclude difficult to reach groups (such as people living in remote areas). The small sample size reduces the statistical precision of the analysis and often makes it impossible to compare different population groups (male and female-headed households, age groups, kinds of agricultural production, etc.).

Contextual analysis: Program outcomes are affected by economic, political, ecological and other factors in the areas surrounding the project. Similar projects may have very different outcomes depending on these contextual factors, but the cost of data collection usually limits a systematic assessment of these factors.

Addressing complexity: While evaluators and other stakeholders frequently state that their programs are complex⁵, very few evaluations are able to address complexity in a systematic way. This is partly due to the large amounts of data required for a complexity analysis, but also because it will usually be necessary to incorporate systems analysis – which often requires the use of dynamic modeling and real-time data, which usually cannot be generated through conventional evaluation methods. The evaluation of the SDGs is an example of an evaluation that ideally requires a complexity focus and the use of systems analysis, but which are usually not possible.

Sustainability and longitudinal data: Most programs are designed to continue to operate, delivering services over a relatively long time horizon. However, for budget and organizational reasons, most evaluations only cover the period during which a program is being implemented and perhaps one or two years after it becomes operational. So, in most cases almost no information is obtained on whether the program and its benefits were sustainable. Furthermore, very few evaluations are able to collect information on the years before the project begins.

We will argue that the new information technologies that are becoming available under the 4IR offer exciting new ways to address these issues.

While all of the above challenges will continue, many 4IR development programs will face new evaluation challenges, including:

- Many programs are becoming larger and more complex.
- Programs will be more dynamic and rapidly changing.
- The private sector will be recognized as a major development partner and public-private partnerships will become more central to development.
- Increasingly large volumes of data will be generated more rapidly, from
multiple sources and of many different kinds (e.g. satellite and drone images, numerical, audio-visual, text); these will need to be captured and processed.

- Broad development challenges such as climate change, refugees and unregulated migration.

- The SDGs pursue 17 ambitious, complex and inter-related goals; as discussed above, these provide major challenges for evaluators.

- Increasing demands from governments, donors, civil society and other stakeholders for accountability and for greater access to information on public sector interventions.

- New information technologies are becoming accessible to any organization or individual with access to a computer and a certain level of technical expertise; consequently, governments and donors will no longer have a monopoly on information about their programs.

- Pressures towards social inclusion and popular participation.

Opportunities of the 4IR: Big data and data science offer powerful new tools that can potentially strengthen development evaluation

As well as presenting challenges for evaluation practice, the 4IR also comes with opportunities. Big data is one such opportunity. Big data is often defined as data that is too large to handle on a normal office computer. It is generated very fast, produces huge volumes of multiple kinds of data, and often generates continuous data over long periods of time. Big data will be discussed in more detail in a second article, but the following are some of the exciting benefits for development evaluation in the 4IR era:

- Data is cheap and fast to collect.
- Very large volumes of data can be collected, often covering the total population of interest, not just a small sample.
- Data can be analyzed very rapidly and can provide continuous feedback.
- Evaluators can combine multiple sources of data (surveys, text – including huge PDF files, audio-visual data, satellite and drone images, photographs) into an integrated data platform.
- Evaluators can have much more sophisticated kinds of analysis including predictive modeling, machine learning and artificial intelligence, the integration of multiple data sets, and the analysis of complex programs.
- Analysts can generate continuous records, including for the period before a program begins and after it ends.

Box 2 gives examples of big data sources that could be used to strengthen the evaluation of development programs in the era of the 4IR. The examples go beyond the “disruptive technologies” identified in the AfDB 2019 paper. Potential applications of these techniques will be discussed in the second article.

Challenges for integrating the new big data tools into development evaluation

There are a number of factors likely to affect the speed and extent to which 4IR data science tools and techniques are adopted by evaluators.

Different research frameworks and tools. One important challenge is that evaluators and data scientists use varying frameworks and tools, and consequently many evaluators are not familiar...
The culture of evaluation in the age of big data: The need for a new evaluation paradigm for the 4th Industrial Revolution?

Data scientists make extensive use of real-time data (e.g. posts from twitter and other social media), the intensity of electric light in a community as an indication of the level of economic activity), while evaluators are more familiar with data generated from surveys, project records and statistics. As a result, many evaluators are concerned about data scientists’ perceived lack of concern about issues relating to bias, data quality and construct validity.

There are differences with respect to the role of theory. While many evaluators consider theory-based evaluation (e.g. theory of change) to be critical in the design of an evaluation and the interpretation of the findings, many data scientists either question the need for theory in the analysis of big data or they have a different approach.

One area of debate concerns the difference between causality and correlation. While the use of comparison group designs is a central part of many kinds of evaluation methodology, data scientists often work with large correlation matrices, sometimes arguing that with a large enough number of observations and correlations it is not necessary to discuss attribution and causality.

Another area of debate concerns data scientists who claim that they...
have complete population coverage whereas evaluators normally work with relatively small samples. Evaluators reply that data science is rarely able to achieve complete population coverage and that the groups left out tend to have important characteristics (such as being poorer) so that these exclusions may result in serious bias. Evaluators also argue that data scientists are often not concerned with bias and do not have a systematic approach to estimate exclusions and their impact.

Weak institutional linkages between data centers and evaluation offices. Based on a limited number of interviews with large development agencies, it seems that the linkages between data centers and evaluation offices are often not well established and many data analysts do not see support for evaluation as one of their primary areas of activity. Similarly, many evaluation offices have not taken the initiative to develop relations with a data center.

Limitations on access to big data. Another practical challenge is that many organizations, particularly those that are smaller and less well funded, may only have limited access to many kinds of big data.

Limitations of integrated capacity development programs. Many evaluation training programs do not include data science methodology, and similarly many data scientists have never received training in evaluative thinking and methods.

Adapting development evaluation in a world of 4IR – rethinking the culture of evaluation.

The era of the 4IR offers powerful new tools for the collection (e.g. big data, satellites and drones, IoT and blockchains) and analysis (e.g. predictive analytics and artificial intelligence) of the increasingly complex development programs that are likely to evolve. These tools make it possible for evaluation to work with larger and more disaggregated data, to broaden the geographical coverage to include contextual analysis, to incorporate vulnerable and difficult to reach groups, and to use systems analysis and other tools to analyze programs operating in complex environments. It also becomes easier to study behavioral change, attitudes and group dynamics.

At the same time, interventions to be evaluated are likely to become larger, multidimensional, interconnected, dynamic and constantly changing. These interventions will also affect and be affected by global challenges such as climate change, growing refugee problems, complex migration patterns and human trafficking. The powerful resources mobilized by 4IR actors also mean that multiple unintended outcomes must also be evaluated, as well as the ‘dark side’ of the 4IR such as fake news and new ways to manipulate consumers and voters and to create social divisions. The sources of potential biases built into remote decision-making algorithms must also be assessed.

For the reasons discussed above, the current evaluation culture and the most common evaluation methodologies are not well suited to fully understand and evaluate this new world. As a result, we are left with the question as to whether the different factors that make evaluation resistant to change can be overcome and allow a new evaluation culture to evolve. Two alternative future scenarios for development evaluation can be imagined. In the utopian scenario, evaluation stakeholders and evaluators will adapt to take advantage of the powerful new tools that are becoming available, and a new and vibrant evaluation culture will evolve that will become a powerful player in guiding the evolution of development programs. There is also a dystopian
(pessimistic) scenario in which evaluation culture is resistant to change, does not take advantage of the new resources and becomes increasingly marginalized and replaced by new 4IR management decision-making tools.

In conclusion, we argue that there is an urgent need to completely rethink current evaluation methodologies to take full advantage of the new information technologies available. This allows evaluators to prepare evaluation agencies and address the challenges of assessing the more complex development strategies and programs that are already evolving. The opportunity is there to adapt the current evaluation culture(s) to the demands of the 4IR world. Will the evaluation community seize the opportunity?

Coming in Evaluation Matters Q2 2020: Part 2 - Transforming evaluation in the 4th Industrial Revolution: Exciting opportunities and new challenges

Part 2 of this article, coauthored by Michael Bamberger and Pete York, will appear in the 2nd Quarter 2020 edition of Evaluation Matters. The article will discuss how evaluation practice will be affected by the five “disruptive elements” of the 4IR, the exciting opportunities that these offer for evaluation, as well as the methodological, organizational and political challenges they will bring. It will also review cutting-edge evaluation technologies, which are already evolving in industrialized nations that will become available throughout Africa in this new era; and will look at the disruptive challenges that evaluators will be faced with during the process of transformation. One or more case studies will illustrate how the new evaluation approaches are being, or could be, applied to the different kinds of evaluations that the African Development Bank conducts.
The culture of evaluation in the age of big data: The need for a new evaluation paradigm for the 4th Industrial Revolution?

References


Endnotes

1. A number of authors have also identified the following reasons for conducting an evaluation: (1) assessing the broader development impacts of a program (for example, contribution to one or more of the SDGs); (2) adapting an intervention to a new context or to a change in the policy environment; (3) contributing to resource allocation decisions on alternative programs; (4) helping to identify emerging problems and building consensus on the causes of a problem and how to respond; (5) supporting public sector reform and innovation; and (6) contributing to global development initiatives (MDGs) and challenges (climate change, human trafficking). Source: Bamberger and Mabry 2020 combining ideas from various authors.

2. An industrial revolution is a fundamental technological transformation that transforms the rest of the social and economic system and where there is a strong interdependence of the different elements of the system. Economic historians have identified three industrial revolutions: The First Industrial Revolution (1760-1840); The Technological revolution (Second Industrial Revolution : 1840-1920); and The Digital Revolution (The Third Industrial Revolution: 1975-2020). The fourth industrial revolution, which does not yet have a universally accepted name, is based on dramatic changes in information technology and how it is applied and involves the blurring of boundaries between the physical, digital and biological worlds.

3. The complexity literature identifies at least four complexity dimensions that affect program design, implementation and outcomes: (1) the complexity of the program itself, (2) complex interactions among multiple stakeholders, (3) the systems within which the program operates and (4) the nature of non-linear causality (Bamberger, Vaessen and Raimondo 2016 Dealing with complexity in development evaluation).

4. in Evaluation Matters Q2 2020 edition, forthcoming
Michael Bamberger has a Ph.D. in Sociology from the London School of Economics. He has been involved for over forty years in the evaluation of development programs in Africa, Asia and Latin America, including 25 years as a senior sociologist with the World Bank. His work has focused on poverty and social exclusion, gender equality, urban development and the challenges of evaluating the MDGs. Over the past few years he has worked on the opportunities and challenges for integrating new information technology into the evaluation of development programs. His recent publications include: “Evaluation in the age of big data”; [with Jos Vaessen and Estelle Raimondo] “Dealing with complexity in development evaluation”, [with Linda Mabry] “RealWorld Evaluation: working under budget, time, data and political constraints”, and “Evaluating the Sustainable Development Goals through equity-focused and gender-responsive evaluations.” Over the past 20 years he has consulted with 10 UN agencies, multilateral and bilateral development agencies, development banks, foundations, NGOs, and governments on the evaluation of development policies and programs.
2019 in review

In 2019, IDEV conducted and facilitated evaluations, knowledge sharing events and capacity development activities in more than 26 countries. Our work not only helped strengthen the efforts of the AfDB, but also contributed to those of the whole evaluation community, including governments and parliamentarians, by providing an opportunity to learn from experience, and to plan and deliver development results to the highest possible standards.

In 2019, the IDEV team delivered 12 evaluations and an evaluation manual.

Evaluations delivered included:

- Two Project Completion Report validation syntheses (Synthesis Reports on the validation of 2016 and 2017 Project Completion Reports).
- A project cluster evaluation (Strengthening Agricultural Water Management to Feed Africa).
- Two country strategy and program evaluations (AfDB Country Strategy and Program Evaluation – Kingdom of Eswatini, South Africa).

IDEV Achievements in 2019

- 13 Evaluation products
- 8 Knowledge sharing events
- 12 Evaluation capacity development activities
Four corporate evaluations

1. Independent Evaluation of the Implementation of the Development and Business Delivery Model of the AfDB.

2. Evaluation of the Middle-Income Countries Technical Assistance Fund.


A regional integration strategy evaluation (West Africa RISP Completion Report Validation).

One thematic evaluation (Evaluation of the Bank’s utilization of the Public Private Partnership Mechanism), and


Eight knowledge-sharing events were organized, in addition to the publication of four editions of Evaluation Matters. In 2019, IDEV also completed a comparative review of sanctions practices across five Multilateral Development Banks which resulted in the publication of a knowledge product.

12 evaluation capacity development activities were organized in 2019, in partnership with the Evaluation Platform for Regional African Development Institutions (EPRADI), the Centers for Learning on Evaluation and Results (CLEAR) and the Twende Mbele peer learning partnership. This includes six webinars that were organized to promote evaluation capacity development and share lessons from IDEV evaluations at the AfDB and in Regional Member Countries.

IDEV organizes a capitalization workshop on the evaluation synthesis of gender mainstreaming at the African Development Bank

On 24 January 2020, a capitalization workshop on gender mainstreaming was co-organized at the African Development Bank Headquarters in Abidjan by IDEV and the Bank’s Gender, Women and Civil Society Department. It aimed to share the knowledge gained from the recent Evaluation Synthesis of Gender Mainstreaming at the AfDB with gender practitioners and champions at the Bank so as to invoke discourse on the implications for the new Bank Gender Strategy currently being developed.

More than 45 Bank staff from various divisions and backgrounds, including from the Bank’s country and regional offices (connecting through video conference), shared their insights about the evaluation’s findings and lessons based on their experience, and at the same time, they identified which aspects are important for the next Gender Strategy of the Bank.

Find out more:

Have you checked out the AfDB Evaluation Results Database (EVRD)?

evrd.afdb.org

Access over 1,030 evaluation documents, 3,850 recommendations and 3,550 lessons learned in one central place

Filter your results by report type, sector, region and country

Each evaluation in the EVRD provides a quick overview of the program, project, or policy under scrutiny; the evaluation findings and recommendations as well as the link to the full report.
Lessons on the Effectiveness of Development Partnerships: An Evaluation Synthesis

Given the importance of partnerships for the Bank, IDEV is for the first time undertaking a comprehensive thematic evaluation of the Bank’s partnerships. This Evaluation Synthesis on Partnerships constitutes the first phase. The purpose of the evaluation synthesis is to facilitate learning, both within and outside the AfDB, by broadly assessing the knowledge accumulated in recent years on the development effectiveness of partnerships and donor coordination activities. It brings together evaluative evidence and research from 38 studies on partnerships at the global, corporate, country and project levels in International Financial Institutions with similar partnership objectives and institutional processes as the AfDB.

The evaluation synthesis aimed to: (i) generate an overall picture of the performance of development partnerships (DPs); (ii) identify lessons relevant to the improvement of DPs going forward; and (iii) inform the ongoing comprehensive evaluation of the AfDB’s partnerships.

For future DPs, the synthesis has identified the following key lessons:

- DPs work best when partners are adaptive and able to respond to partnership opportunities arising from a client’s needs and priorities.
- It is important to strengthen organizational capacity to deliver on partnerships’ objectives;
- Partnerships that are mainstreamed are more likely to realize their full potential;
- Selectivity and good management contribute to value addition and the attractiveness of an organization for new partnerships;
- Weak monitoring & evaluation systems, and reporting requirements need to be improved to ensure the performance of DPs.

Find out more:

Fourth Quarter 2019: Made in Africa Evaluations Volume 2

In the call for contributions for this edition on Made in Africa Evaluations (MAE), we received many articles, thus warranting two volumes. The first volume focused on the theoretical approaches of MAE. This volume highlights some of the practical applications. Much of the discourse to date on MAE has focused on the conceptual aspects. Therefore, this volume aims to take the discussion a little further in understanding what it is, by providing practical experiences on what the concept looks like in practice.


Third Quarter 2019: Made in Africa Evaluations Volume 1

Development evaluation approaches have grown into a largely uniform global practice, in particular among development organizations ascribing to internationally agreed norms such as the OECD-DAC evaluation criteria. Some evaluation practitioners have called into question the relevance and effectiveness of current evaluation approaches in the African context, calling for a “Made in Africa” evaluation that takes into account local values, assumptions, and practices. This edition takes stock of some of the theoretical approaches towards a “Made in Africa” evaluation, exploring indigenous approaches and how they could fast-track the achievement of the continental development agenda.


Second Quarter 2019: Best practices and innovation in evaluation

The field of evaluation is on the move – in tracking progress on Agenda 2030, dealing with increasingly complex development interventions, new technologies and sources of data, and more sophisticated evaluation methods. Sharing good practices and innovations in evaluation can help evaluators to learn from each other, to tackle challenges and continually strengthen the profession. This edition of Evaluation Matters aims to showcase selected good, new or innovative evaluation methods that have contributed to better evaluations of development interventions, with a view to improving project/program planning, design and implementation.


First Quarter 2019: Gender in Evaluation Volume 2

Women continue to suffer significant economic, political, legal, social and cultural disadvantages in almost all societies. Evaluations of projects, programs and policies must take into account these disadvantages and provide stakeholders with sound and compelling evidence to better inform the planning and implementation of future development interventions. This edition complements Evaluation Matters Quarter 4 2018 by providing examples of how selected individuals and institutions have been able to concretely integrate Gender Responsive Evaluation approaches into their work.
