How can we Strengthen National Evaluation Systems?
—Perspectives from around the globe

“‘It’s a bit like the Titanic: Do you want to course correct and miss the iceberg, or do you want to set up a comprehensive review of why it hit the iceberg and what lessons you could learn?’”

—Trevor Davies in How to Speed up Global Development’s Learning Cycle, page 92
A QUARTERLY KNOWLEDGE PUBLICATION OF THE AFRICAN DEVELOPMENT BANK
September 2013 . Volume 2 . Number 3

3 From the Director’s iPad

Rakesh Nangia

5 Defining Capacity and Evaluation Capacity Development

6 Towards a Shared Framework For National Evaluation Capacity Development

Marco Segone, Caroline Heider, Riitta Oksanen, Soma de Silva, and Belen Sanz

28 Perspectives from Around the Globe

29 Africa

- Strengthening National Evaluation Systems
  Samuel Kouakou, Côte d’Ivoire

- The importance of Strengthening National Evaluation Systems
  F. Kelemework Fasika, Ethiopia

- Developing a National Evaluation System in South Africa,
  Ian Goldman, South Africa

- Patterns and Influences in the Supply and Demand of Evaluation and Monitoring in Uganda’s Public Sector over the Past Two Decades
  David Rider Smith, Uganda

- Strengthening National Evaluation Systems
  El Hadji Gueye, Senegal

- National Evaluation Capacity Development: Past and Future Directions
  Mohamed Manai, AfDB

66 Latin America

- Speaking Truth to Power: The Role of National Evaluation Systems in Democratic Governance
  Claudia Maldonado

76 North America

- Rebirth, Restoration, Reclamation, and Responsibilities of the Evaluation Function of Africa,
  Sukai Prom-Jackson, UNDP Independent Evaluation Office

91 Food for Thought

- How to Speed up Global Development’s Learning Cycle
  Trevor Davies, KPMG

- Developing Evaluation Capacity, but Which Capacity?
  Samer Hachem, AfDB

- Framework for Understanding NECD Definitions and Terminology

- Towards an Evaluation Culture in the AfDB:
  Practices and Challenges,
  Mohamed Manai, AfDB

- Checklist for developing a national evaluation system
FISCAL PRESSURES, GREATLY magnified by the global financial crisis, brought not only pressure on official development assistance support but also increased demands from external stakeholders. The focus on results is unprecedented, and goes hand-in-hand with the demand for useable, results-based monitoring and evaluation (M&E) systems to provide tangible evidence of outcomes, as well as support for evidence-based decisions.

The traditional role for the independent evaluation function housed in various multilateral development banks is to focus on the performance of programs financed by the respective institutions. Most of the evaluations conducted are ex-post. However, regardless of how well they are conducted and how influential they are with respect to reforms, they are a small part of the national public investment program.

It is obvious that influencing the entire investment program is the smarter way to go. In order to do so, development institutions, including their respective independent evaluation departments, need to rethink their traditional role. Building capacity in M&E has been part, albeit a small part, of the program of these departments for many years. But capacity is only one element of the system, which includes training staff; establishing appropriate organizational structures; ensuring good-quality data and flow; and, of course, using the evidence to inform decisions. What role should the evaluation departments play in building these systems? Is it even within their mandates?

I would submit that continued capacity-building efforts in the absence of a strengthening of the remaining elements of the overall system is a wasted effort. The value of a
holistic perspective is one of the lessons that various evaluation reports (on capacity building) have highlighted over the years. Yet we, the evaluation community that constantly advocates for learning lessons from experience, seem to ignore this advice in our own capacity-building initiatives.

Ideally, we should begin with an analysis of the existing M&E system, including identifying the gaps. This should, of course, be done with counterparts in governments, which one would hope includes some champions.

All of this is expected to influence the public investment program, and thus a performance- and results-based budget allocation. It is clearly very early days, but the overarching vision and direction are laudable. Maybe, just maybe, the evaluation community will learn from the experience and play a catalytic role in South-South learning and scaling up across the developing globe.

Rakesh Nangia is the Director of the Operations Evaluation Department of the African Development Bank. Prior to joining the AfDB, he spent 25 years at the World Bank, where he held several positions including Director of Strategy and Operations for the Human Development Network and Acting Vice-President for the World Bank Institute.

He attended the Indian Institute of Technology in Delhi and Harvard University and holds degrees in business administration and engineering.
Defining Capacity and Evaluation Capacity Development

THE UNITED NATIONS Development Group defines capacity as the ability of individuals, institutions and societies to perform functions, solve problems, and set and achieve objectives in a sustainable manner, and capacity development as the process through which the abilities to do so are obtained, strengthened, adapted and maintained over time.

For OECD, evaluation capacities are the ability of people and organizations to define and achieve their evaluation objectives (OECD, 2006). Capacity involves three interdependent levels: individual, organizational and the enabling environment. These interrelated capacities function together to demand, supply and use evaluation. Evaluation capacities include the power to set the evaluation agenda and to determine what is evaluated and what questions are asked (OECD, 2011). Further, capacities include the ability not just to produce evaluations but also to manage evaluation processes and effectively use evaluation results to influence policy and programme decisions. It is useful to distinguish between the capacity to manage evaluations and the capacity to conduct them, as both are necessary (Feinstein, 2009).

Evaluation Capacity Development (ECD) is understood as the process of unleashing, strengthening and maintaining evaluation capacities.

Putting evaluation to work
Introduction

National ownership and leadership as overarching factors for ensuring relevant development outcomes

The Paris Declaration on Aid Effectiveness endorsed in 2005, and the 2008 Accra follow-up meeting, state that national ownership and leadership are overarching factors for ensuring good development outcomes. The implication for the evaluation function is fundamental.

The principle of ownership means that partner countries should own and lead their own country-led evaluation systems, while donors and international organizations should support sustainable national evaluation capacity development. The 2012 Busan High-level forum reaffirmed the above principles, while recognizing that the international development arena has changed significantly. As a result, new modalities such as south/south and triangular cooperation, and new stakeholders such as Civil Society Organizations (CSOs), have been explicitly recognized in the Busan’s “Partnership for effective development cooperation”. The Busan principles are also supported by the United Nations. The UN General Assembly has requested the UN system to pursue and intensify its efforts.

This paper was published by EvalPartners in the book Evaluation and Civil Society. Reprinted here by permission.
to strengthen evaluation capacities in program countries, taking into account national conditions and ensuring respect for national ownership, strategies and sovereignty.

**National evaluation capacities as part of good governance**

As countries take greater ownership of, and leadership in, their development processes, they have also increasingly developed their systems to lead, manage and account for resources invested in these processes and results produced with them. Results-based budgeting and management have been part of the agenda to strengthen governance and go hand in hand with capacities for monitoring and evaluation. Evaluation capacities empower stakeholders—from national governments to CSOs—to question, understand and take charge of necessary changes to development processes as and when evidence indicates policies, programs and projects can be more successful, effective and efficient. Exercising evaluation in an independent, credible and useful way is essential to realize the contribution it can make to good governance, including accountability from governments to their citizens and their development partners, transparency in the use of resources and their results, and in learning from experience. Understanding how powerful evaluation can be in making informed decisions will lead to the creation of an enabling environment and demand for evaluation as an integral part of debates and decision-making processes, and will stimulate development or strengthening of evaluation systems to meet these demands.

**Strengthening existing national evaluation capacities**

In a number of countries, public policy evaluation functions are in place (UNEG, 2011). Examples are those set up through the implementation of Poverty Reduction Strategy Papers, which link monitoring and evaluation systems to poverty reduction. Seventy per cent of countries in Africa have this type of set up with different institutions having the mandate for evaluation. There is a positive move towards independence of the evaluation function in several countries. A growing number of countries have included the evaluation function in their Constitution or mandated it via Acts of Parliament. In some countries the involvement of civil society is gaining momentum through dialogue and peer review mechanisms.

However, while many governments have defined evaluation functions, and put in place institutional arrangements for evaluation to deliver evidence to inform policy-making, existing capacity varies very significantly from country to country.
This means that any initiative to strengthen national evaluation capacities should recognize existing capacities and strengthen them, being sensitive to different levels of capacity, rather than stimulating new and parallel capacities.

**The growing role of civil society organizations in national evaluation capacity development**

Along the lines of the recent Busan Declaration, CSOs can and should play a central role in advocating for transparency in the allocation and expenditure of public budgets; accountability for the implementation of public policies; strengthening the demand and use of evaluation to inform evidence-based policy-making; and, strengthening capacities of qualified evaluators to produce credible and useful evaluations based on national and international evaluation standards.

While national evaluation capacity is often understood as government capacity, it should be a country-based capacity, including Voluntary Organizations for Professional Evaluation (VOPEs), universities, think tanks and Non-governmental Organizations (NGOs). CSOs and parliamentarians need evaluations to support their understanding of issues and participation in decision-making. This can strengthen the quality of democracy whereby informed citizens are able to influence decision-making.

In the last decades, members of both civil society and the private sector have been playing increasingly central and active roles in promoting greater accountability for public actions through evaluation.

National and regional VOPEs grew from 15 in the 1990s to more than 150 nowadays.

In this context, 25 organizations launched EvalPartners, an international collaborative initiative to contribute to the enhancement of the capacities of CSOs—notably VOPEs—to influence policy-makers, public opinion and other key stakeholders so that public policies are evidence-based, equitable and effective. The main objective is to enhance the capacities of CSOs/VOPEs to engage in a strategic and meaningful manner in national evaluation processes, to be able to influence country-led evaluation systems.

**Towards a shared framework for national evaluation capacity development**

In addition to governments, VOPEs and NGOs at country level, a multitude of stakeholders is currently engaged in supporting National Evaluation Capacity Development:

> Since a number of different names are used to describe these groups, the term Voluntary Organizations for Professional Evaluation (VOPEs) has been introduced. VOPEs include formally constituted associations or societies, as well as informal networks and communities of practice. Their memberships are open not only to those who conduct evaluations but also to those who commission and utilize evaluations and those engaged in building the evaluation field.

> The United Nations Evaluation Group (UNEG) and its members, the Evaluation Cooperation Group (ECG) of the International Financial Institutions (IFIs), the Development Assistance Committee (DAC) of the Organisation for Economic Cooperation and Development

> IOCE maintains a database of VOPEs on its www.IOCE.net website, including an interactive world map showing the locations of and contact information for national and regional VOPEs.

> Please visit http://www.mymande.org/evalpartners for additional information.
(OECD), the regional CLEAR centers, among others. However, in certain cases different actors have not coordinated their support and initiatives, making it difficult to create positive synergies and efficiency. In extreme cases, duplications and parallel processes are supported by different agencies. Therefore, there is a need to create synergies based on a shared framework for National Evaluation Capacity Development. This will help guide development activities in a comprehensive way to strengthen national evaluation systems as a whole, in addition to providing guidance on good practice, based on evidence of what works and why.

Defining Capacity and Evaluation Capacity Development

The United Nations Development Group (UNDG) defines capacity as the ability of individuals, institutions and societies to perform functions, solve problems, and set and achieve objectives in a sustainable manner, and capacity development as the process through which the abilities to do so are obtained, strengthened, adapted and maintained over time.

For OECD, evaluation capacities are the ability of people and organisations to define and achieve their evaluation objectives (OECD, 2006). Capacity involves three interdependent levels: individual, organisational and the enabling environment. These interrelated capacities function together to demand, supply and use evaluation. Evaluation capacities include the power to set the evaluation agenda and to determine what is evaluated and what questions are asked (OECD, 2011). Further, capacities include the ability not just to produce evaluations but also to manage evaluation processes and effectively use evaluation results to influence policy and program decisions. It is useful to distinguish between the capacity to manage evaluations and the capacity to conduct them, as both are necessary (Feinstein, 2009).

Evaluation Capacity Development (ECD) is understood as the process of unleashing, strengthening and maintaining evaluation capacities. Capacity development is a long-term, endogenous change process that takes place in the context of ongoing partner and donor efforts to strengthen related systems of management, governance, accountability and learning, to improve development effectiveness. The best capacity development approaches are flexible, adapted and sustainable.

Strengthening evaluation capacities is not an end goal in itself, but should be seen, rather,
as a means to support more effective development activities and informed policy-making. Evaluation implicates actors across the entire range of stakeholders involved in and affected by public policy: governments, intended beneficiaries, civil society, and the general public.

A Systems Approach to National Evaluation Capacity Development

National Evaluation Capacity Development (NECD) is a complex field in which several stakeholders have different roles to play based on their respective value added. This complexity encourages the use of a systems approach to NECD. This means that it is necessary not only to look at actors at different levels and across sectors but also, crucially, to look at the network of relationships or connections between them. Such a viewpoint illustrates the fact that weaknesses in capacity at any level or with any key actor will affect the capacity of the whole system to deal with a problem in order to achieve a goal. Therefore, a systems approach to NECD is needed.

Individual and institutional evaluation capacities enabled by a supportive environment

In the past, evaluation capacity development focused on strengthening the capacities of individuals’ knowledge and skills. However, it is by now clear that capacity development should be based on a systemic approach that takes into account three major levels (individual, institutional, and external enabling environment), and two components (demand and supply) tailored to the specific context of each country.

Figure 1: A systemic and integrated approach to national evaluation capacities development

---

5 “Supply” refers to the capability of professional evaluators to provide sound and trustworthy evaluative evidence; “Demand” refers to the capability by policy makers and senior managers to request sound and trustworthy evaluative evidence with the aim of using it in strategic decision-making processes.
Enabling Environment: the enabling environment provides a context that fosters (or hinders) the performance and results of individuals and organizations

- Strong evaluation culture:
  - evaluative (critical) thinking;
  - protective environment;
  - managers and other decision-makers value and use evaluation.
- Public administration committed to transparency and managing for results and accountability, through results-based public budgeting and evidence-based policymaking.
- Political will to institutionalize evaluation.
- Existence of adequate information and statistical systems.
- Legislation and/or policies to institutionalize monitoring and evaluation systems.
- Duty bearers, i.e. Governments and national authorities, have capacity and willingness to be accountable for results.
- Rights holders i.e. citizens and CSOs, have capacity to demand that policy and programs are monitored and evaluated.
- National VOPE exists, has the necessary capacities, is recognized and given a consultative role.
- National evaluation standards and norms developed and implemented.

Institutional Level: the institutional framework in which individuals work needs to provide a system and structures to perform and attain results individually and collectively, as an organization

- Evaluation policy exists and is implemented.
- An evaluation unit with a clearly defined role, responsibilities and an independent place in the institutional structure exists and is functional.
- Quality Assurance system exists and is functional.
- Independence of funding for evaluations.
- Adequate (number & expertise) staffing of the evaluation function.
- System to plan, undertake and report evaluation findings in an independent, credible and useful way exists.
- System to establish a formal management response to the recommendations of evaluations feeding back to policies and programmes, follow-up of implementation, exists.

Towards a shared framework for National Evaluation Capacity Development

- Open dissemination of evaluation results.
- Knowledge management systems in support of the evaluation function exists and is used.
- Technical competence for conducting evaluations exist.
Individual Level: the individual whose knowledge, skills and competences are essential to perform tasks and manage processes and relationships

- Senior management capacity to:
  - strategically plan evaluations;
  - assure the development of relevant and appropriate Terms of Reference for evaluations;
  - manage evaluation for independence and credibility;
  - promote the use of evaluation findings;
  - follow up recommendations.
- At mid-management level, understanding of the role of evaluation as a tool for effectively achieving development results.
- Identify and support leaders or natural champions.
- Behavioral independence and professional competences of those who manage and/ or conduct evaluations.
- Promote capacity development and involvement of a growing number of evaluators.

Source: Adapted from Segone, M. (2010), Moving from policies to results by developing national capacities for countryled monitoring and evaluation systems.

The enabling environment for evaluation is determined by a culture of learning and accountability, meaning the degree to which information is sought about past performance, the extent to which there is a drive to continuously improve, and to be responsible or accountable for actions taken, resources spent, and results achieved. Such culture is embedded in tacit norms of behavior, the understanding of what can and should—or should not—be done, in many cases behaviors being role-modeled by leaders. Through a set of values and attitudes supporting evaluative (critical) thinking within an organization, individuals are more self-directed learners and use information to act; to take higher risks but, also to develop a greater sense of personal accountability and responsibility; and, to consult, coach, and support each other more. In this context, organizations with a culture of evaluation are able to develop innovative ideas and strategies; change more quickly according to variations in the external environment; and, increase efficiency and effectiveness by systematically using lessons learned to improve programs and policies. There is less direction from top management and a much more positive attitude and self-accountability at all organizational levels. An organization with a culture of evaluation has an effective, structured and accepted use of evaluation to support change and development. Managers value and use evaluation findings to test out innovation or assess progress towards expected results. These tacit norms of behaviors are, or should be codified in government legislation and /or an evaluation policy that expresses the leadership's or an organization's commitment to learning, accountability and the evaluation principles. A two-tier strategy should be put in place to strengthen the capacity of dutybearers (policy-makers) to demand and use sound evidence while developing rights-holders' (i.e. citizens, CSOs, etc.) capacity to demand and to assess policy implementation, by putting in place systems and mechanisms to engage citizen groups, and to capture and utilize their feedback.
An enabling environment is also supported or created through governance structures that demand independent evaluation, be it through parliaments or governing bodies, and is further enhanced through VOPEs that set standards and strive towards greater professionalism in evaluation. Therefore, VOPEs should be supported to enable them to foster indigenous demand and supply of evaluation, including by setting national evaluation standards and norms. There are also examples of governments soliciting the advice and involvement of VOPEs in not only the formulation of evaluation policies and systems, but also in the implementation of evaluations consistent with those policies (Presidency of South Africa, 2011).

The structural independence of an evaluation function is important to set an enabling environment: by placing the evaluation function in such a way that it is not controlled by the person or function responsible for the policies, strategies or operations that are evaluated, and in a way that creates an environment that enables greater independence, credibility and utility. Ideally, the enabling environment is such that decision-makers proactively demand impartial evaluations to inform their debates and choices, which increases the usefulness of evaluations.

The institutional framework for evaluation ensures that a system exists to implement and safeguard the independence, credibility and utility of evaluation within an organization. Such an institutional framework has the following characteristics:

- Includes a system of peer review or assurance that the evaluation function is set up to safeguard and implement the principles of independence, credibility and utility.

- Establishes safeguards to protect individual evaluators—evaluators, evaluation managers, and heads of evaluation functions—when exercising their independence, including transparent and credible processes for the selection, appointment, renewal (if applicable) and termination of contract of the head of evaluation and assurances that evaluation staff do not suffer in promotion exercises.

- Ensures a multidisciplinary team exists in the evaluation function, or on an evaluation team, that ensures credibility of evaluation by understanding multiple dimensions of evaluation subjects and combining the necessary technical competence.

- Secures the independence of funding of evaluations, at an adequate level, to ensure that necessary evaluations are carried out and that budget holders do not exercise inappropriate influence or control over what is evaluated and how. Therefore, funding should be under the direct control of the head of the evaluation function, and should be adequate for a reasonable work program. The adequacy of funding for evaluation can be determined through a ratio of evaluation resources (human and financial) over the total program size (some organizations aim to allocate 1 to 3 per cent of their total budget for evaluation) or the coverage rate of evaluations (for instance, evaluating 25 per cent of the overall program).

- Combines measures for impartial or purposeful selection of evaluation subjects to ensure impartiality, on the one hand, and increased utility, on the other, by making deliberate choices linked to decision-making processes.
To ensure utility, for instance of strategic or thematic evaluations, these may be chosen purposively (rather than to establish a representative sample) to link the conduct of the evaluation to the information needs and decision-making processes of those using evaluation findings. In these cases, a consultation process to determine what the most important/strategic topics are for evaluation is important to ensure various stakeholder needs are considered.

- **Sets out a system to plan, undertake and report evaluation findings in an independent, credible and useful way.** To increase objectivity in the planning and conduct of evaluation, systems are needed that increase the rigor, transparency and predictability of evaluation processes and products. Such systems can include more or less detailed process descriptions or guidelines for the design of evaluations, for preparatory work that needs to be conducted and reported on, and for reporting findings. The processes should have built-in steps for communication, consultation and quality assurance, and be communicated to stakeholders to enhance transparency and to secure their willingness to share information.

- **Institutes measures that increase the usefulness of evaluations,** including the sharing of findings and lessons that can be applied to other subjects. Evaluations should be undertaken with the intention—of stakeholders and evaluators—to use their results, and management’s responses developed and implemented. The timeliness of planning and conducting evaluations and presenting their findings is equally important to ensure utility. Finally, accessibility means that evaluations are available to the public, can be retrieved (for instance through a user-friendly website), are written in ways that are understandable (clear language and limited jargon), and are distributed to a wide group of relevant stakeholders in formats appropriate to different audiences.

Even with structures and systems in place, the independence and impartiality of evaluation depends on the integrity and professionalism of individuals as evaluators, evaluation managers, and the head of evaluation. The profession, and its ethics, requires limiting personal biases to the extent possible. In some circumstances, external evaluators are believed to exercise greater independence than those who work in an organization, because they are less exposed to institutional or peer pressures or have not absorbed the corporate culture in an unquestioning way. However, individual or intellectual independence is dependent on the individual, who demonstrates in his/her behavior the adherence to, and practice of, the evaluation principles: avoiding conflict of interest; acting with integrity and with an independence of mind; engaging in evaluations for which they are competent; acting impartially; and, undertaking an evaluation with a clear understanding of the clients or decisionmaking process and how these need to be informed (see AEA, 2003). Many VOPEs and evaluation functions of national and international organizations have adopted codes of conduct for evaluators. In addition to these, the UN Evaluation Group developed ethical guidelines (UNEG, 2005). Debates about professional standards and the accreditation of evaluators and evaluation managers have been ongoing for many years within professional forums.
**Development of national evaluation capacity**

An institutional framework to institutionalize the evaluation function should be developed within the organization. This should include assistance to develop an evaluation policy, which foresees the setting-up or strengthening of a credible evaluation department, including endorsement of evaluation standards and quality assurance systems (including peer review). A diagnosis of the existing evaluation function and/or system (or, if not existing, of the preconditions to develop it), which would include functional clarity, effective human and financial resources management and a robust coordination mechanism, should be carried out to enable the crafting of a context-specific evaluation policy and system. The diagnostic would identify bottlenecks in the policy or system which could be addressed through targeted technical support with the intention of strengthening country capacity.

A good practice is to strengthen knowledge management systems in support of the evaluation function. Knowledge has traditionally been fostered at the individual level, mostly through education. However, seen from the perspective of the three levels identified above, knowledge should also be created and shared within an organization through an effective knowledge management system, and supported through an enabling environment of effective educational systems and policies.

**Figure 2: Tips for stakeholders to help reinforce an enabling environment for evaluation**

**Governing bodies (Parliaments, Executive Boards, etc.)**

- Be aware of the threats to independence, credibility and utility of evaluation and demand measures to safeguard these principles.
- Adopt and oversee the implementation of legislation and/or policies, which institutionalize the independence, credibility and utility of evaluation.
- Put in place effective oversight mechanisms over the quality of evaluation.
- Use evaluation findings and recommendations in national and subnational policies, programs and legislation.

**Heads of State, Ministers, Policy-Makers, Chief Executives Officers**

- Demonstrate leadership in setting a culture of learning and accountability.
- Seek and use evidence from evaluations to validate the attainment of goals and objectives and to improve performance whenever possible.
- Understand evaluation as part of good governance that aims to use public resources. Effectively and efficiently to achieve the goals which governments or organizations aim to achieve.
Evaluation Function

- Demonstrate the value added of independent, credible and useful evaluations.
- Raise the awareness of stakeholders of the role and importance of evaluation and of the evaluation principles.
- Contribute to evaluative thinking through awareness building, dialogue, and training.
- Define and build evaluation competencies at different levels of public office.

VOPEs

- Work with governments to set standards as benchmarks which can be used to convince other stakeholders of the importance of the evaluation principles and measures to safeguard them.
- Foster indigenous demand and supply of evaluation.
- Be available to advise commissioners of evaluations on the relevance of Terms of Reference for evaluations, including choices of appropriate designs and methodologies to answer key questions.
- Promote the capacity of evaluators to perform quality, credible and useful evaluations.
- Advocate for equity-focused and gender-sensitive evaluation systems.

Source: Adapted from Heider, 2011, (Conceptual framework for developing evaluation capacities).

At the individual level, a capacity development strategy should strengthen senior management capacity to strategically plan evaluations (and to identify the key evaluation questions); to manage evaluation for independence and credibility; and, to use evaluation.

MacKay (2007) underlines the importance of identifying and supporting leaders or natural champions who have the ability to influence, inspire and motivate others to design and implement effective evaluation systems. Leadership is not necessarily synonymous with a position of authority; it can also be informal and be exercised at many levels. Therefore, the evaluation capacity development strategy should, especially in the initial stages, identify and support as appropriate, national and local leaders in the public administration, in inter-governmental monitoring, and in evaluation groups and national VOPEs.

On the supply side, a capacity development strategy should enhance behavioral independence (independence of mind and integrity; knowledge and respect of evaluation standards; agreed evaluation processes and products) as well as professional competences through formal education; specialized training; professional conferences and meetings; on the job training (such as joint country-led evaluations); and, communities of practice and networking, e.g. VOPEs.
**Figure 3: Tips for stakeholders to help develop an institutional framework for evaluation**

**Governing bodies (Parliaments, Executive Boards, etc.)**

- Be briefed about the evaluation system so as to understand whether the institutional framework includes adequate checks and balances and to become a discerning user of evaluation reports.
- Introduce processes for the selection of the head of evaluation which ensures his/her independence.
- Provide secure, separate and adequate funding for evaluation in the regular program budgets.

**Head of State, Ministers, Policy-makers, Chief executives officers**

- Be briefed about the evaluation system so as to understand whether the institutional framework includes adequate checks and balances, and to become a discerning user of evaluation reports.
- Establish or strengthen evaluation training at graduate and undergraduate levels to provide professional evaluators in appropriate quantities.
- Create professional evaluation positions in support of the evaluation function in government bodies.

**Evaluation function**

- Develop and document systems for the selection, design, conduct and reporting on evaluations.
- Provide briefings on these standards to increase transparency and confidence in the process and products, and so to enhance credibility.
- Develop mechanisms to ensure lessons from evaluation are systematically shared and integrated into debates and decision-making processes.

**VOPEs**

- Provide guidance to set professional standards and good practice standards.
- Assist the government to develop and implement professionalization, accreditation, and credentialing systems.

Source: Adapted from Heider, 2011, (Conceptual framework for developing evaluation capacities).
Fostering demand for and supply of evaluation

A distinction should be made between the capacity of policy-makers/policy advisors to use evidence and the capacity of evaluation professionals to provide sound evidence. While it may be unrealistic for policy-makers/policy advisors to be competent experts in evaluation, it is both reasonable and necessary for such people to be able to understand and use evidence produced by evaluation systems in their policy and practice. Integrating evidence into practice is a central feature of policy-making processes. An increasingly necessary skill for professional policy-makers/policy advisors is to know about the different kinds of evidence available; how to gain access to it; and, how to critically appraise it. Without such knowledge and understanding it is difficult to see how a strong demand for evidence can be established and, hence, how to enhance its practical application.

However, it is also important to take into consideration that the design and implementation of policy reform is a political process, informed by evidence. The use of evidence in policy reform and implementation depends on the combination of capacity to provide quality and trustworthy evidence on the one hand, and the willingness and capacity of policy-makers to use it on the other. The extent to which evidence is used by policy-makers depends, in turn, on the policy environment.

To strengthen an enabling policy environment, policy-makers may need incentives to use evidence. These include mechanisms to increase the ‘pull’ for evidence, such as requiring spending bids to be supported by an analysis of the existing evidence-base, and mechanisms to facilitate evidence-use, such as integrating analytical staff at all stages of the policy implementation.

Civil society organizations, including VOPEs, should play a major role in advocating for the use of evidence in policy implementation. Thinktanks, with the support of mass media, may also make evidence available to citizens, and citizens may demand that policymakers make more use of it.

It is therefore clear that a strategy for developing national evaluation capacities should be tailored to the situation and context in which a specific country finds itself, as illustrated in Figure 4, and detailed below:

- **Vicious circle countries.** Evidence is technically weak and policymakers have little capacity to make use of it. Policy decisions are mainly taken on political judgment and personal experience, which may result in poor policy design and, consequently, poor results. In this case, it is necessary to adopt measures which will simultaneously increase both the demand and supply of evidence, as well as improve the dialogue between producers and users of evidence.

- **Evidence supply-constrained countries.** Although evidence is technically weak, it is increasingly demanded by policy-makers. However, evidence deficiency reduces the quality of decision-making and therefore the quality of services delivered. Policymakers are likely to resent being held to account on the basis of inadequate evidence. Therefore, the priority should be to adopt measures to increase the quantity and quality of evidence. The challenge is to strike a balance between quickly generating improvements to evidence, while laying the foundations for
Figure 4: Capacity development framework addressing the demand as well as the supply side

The quantity and quality of evidence is improving, but it is not demanded for decision-making because policy-makers lack the incentives and/or capacity to utilize it.

- Evidence demand-constrained countries. The quantity and quality of evidence is improving, but it is not demanded for decision-making because policy-makers lack the incentives and/or capacity to utilize it. Policy-makers are likely to be at the very least wary of (or may even actively dislike) having more and better evidence pushed at them when it may not support decisions they have already taken or wish to take. In this case, priority should be given to the adoption of measures to
increase the demand for evidence, as well as to improve the dialogue between producers and users of data.

- **Virtuous circle countries.** Evidence is technically robust and is being used increasingly for decision-making. The production of good (or at least improved) evidence is matched by its widespread (or at least increased) use in decision-making. These two processes are mutually reinforcing, resulting in better policy design and implementation, and ultimately, better development results.

The virtuous circle countries situation is the goal which should be set, since it provides a useful benchmark against which to compare the other three cases. Developing a culture of evidence-based policy-making is a slow process which may take years, but the potential rewards are worth the effort. Essential elements for sustaining a virtuous circle linking evaluation professionals to policymakers are: strengthening the democratic process by requiring transparency and accountability in public sector decision-making, and establishing clear accounting standards and an effective regulatory framework for the private sector.

**Key Principles of National Evaluation Capacity Development**

The number of stakeholders and the range of capacities and capabilities discussed above illustrate the complexities of capacity development and flag the need for planning and implementation tools that are able to address complexity, something that blueprints tend not to be good at. Instead, space needs to be created for analyzing capacity gaps in a participatory way, agreeing on common goals and translating them into a joint strategy. At the same time the flexibility and capability should be maintained to take up opportunities as they arise, learn from experience and change tactics, and work on various levels of capacity at the same time.

The principles of *national ownership and leadership* have to do with the realization that capacity development needs to come from within and/or have strong internal champions, whether it is when developing a national system or a function in an organization. If an idea is imposed from the outside, chances are that capacity development is not likely to succeed. The drive from within relates to the capability “to commit and engage.” It includes:

a) **Ownership**, which manifests itself in having local or internal champions, resource allocations, and engagement. It is, however, difficult to measure, not homogenous across internal stakeholders, and not constant over time;

b) **Leadership**, which is important, especially for setting the culture of learning and accountability, but also must fit with the context and its culture; and

c) **Collective action**, motivation and commitment, meaning that capacity development will not occur if it is one person writing documents, but requires engaging various stakeholders whose capacities and capabilities will be developed.
Capacity development should therefore be underpinned by the fundamental characteristic of national ownership. Taking a capacity development strategy to scale requires linking it to national and local plans, processes, budgets and systems. To be sustained, a comprehensive capacity development response must link to, and draw from, relevant national reforms.

The above is particularly important as capacity development is about transformations and must address how best to manage change within the existing policy environment. The tendency to often look only inside an organization and to downplay the larger institutional context, in which the organization resides, has proven unsuccessful. To ensure continued political commitment and resource support, a capacity development strategy can and often must show both short and long-term gains. Experience shows that capacity traps are more often pertinent to the “soft side” of the policy environment, such as vested interests, ethic and attitudes, rather than to “hard” technical competencies.

The principle of context-specific capacity development means that capacity must be understood in terms of a specific cultural, social and political context. Capacity must be understood as something that exists in degrees at all levels of society: individual, household, community, institutional and national systems. Capacity can exist without outside intervention, though it may be constrained. This implies that one must first understand capacities or elements of capacity that already exist before engaging in any effort to build on or strengthen them.

“Context-specific” capacity development means that no blueprint should be applied, but high degrees of flexibility should be combined with systematic approaches to entail a recognition that capacity development, while systematic, does not (always) follow an ordered process. Instead it includes planned as well as incremental and emergent approaches. On the other hand, when trying to influence culture to become more supportive of learning and accountability, opportunities might arise at short notice, in various settings and without prior notice, and include an informal conversation with a key stakeholder. In these contexts, capacity development initiatives would be incremental and emergent in response to opportunities as they arise. Capacity development also requires recognizing the less tangible aspects of capacities and combining small and large initiatives, depending on context and opportunities. It requires finding the right balance between an operating space that allows capacities to evolve and accountability for capacity development results; a balance that is hard to attain. And finally, capacity development takes time, so it is necessary to stay the course (even through adverse times), but build quick wins into the process (to keep up motivation and build on success), and time to reflect and evaluate whether progress is being made.

The process of developing and agreeing on clear expectations in “country-specific” capacity development is in itself part of the capacity development process. Developing a common understanding of capacity weaknesses through a participatory, structured diagnostic—using tools like the conceptual framework in Figure 1 (p. 11)—can serve as an analytical framework to move attention from resource gaps to recognize broader issues, and a shared vision of the needed capacities are important steps in channelling resources towards clear objectives. Very often
the many actors in capacity development have tacit understandings of what capacity is needed and how its development should happen. Few of them have explicit strategies, which makes it more difficult to find a common ground and work in the same direction. A diagnosis of existing capacities is necessary to find entry points to start capacity development. This diagnosis benefits from being developed in a participatory way that develops ownership at the same time.

Aligning with national evaluation capacity development strategies and systems

Strong national ownership and leadership are the foundation for aligning external support that leads to sustainable results. National evaluation capacity development strategies provide the starting point. When the national strategies are of high quality in terms of having clear priorities, being results-oriented and operational, they are a natural basis for cooperation and alignment. Another important indication of national commitment is the resource allocations in the institutions’ budgets, and staffing for the implementation of the capacity development strategies.

Alignment also means that country systems are used when providing external support. This includes the use of the national institutions’ own systems of planning, financial management, monitoring, reporting and evaluation. Country systems should be fully used when they provide reliable assessments of performance, ensure transparency and accountability of the systems, and take necessary steps to strengthen and reform the systems. Providers of external support may align their contributions to national systems fully or partially. If alignment is only partial, the decisions should be regularly reconsidered with the development of the country systems. Step-by-step progress in alignment means that efforts and results in improving country systems are appropriately acknowledged.

In some cases the biggest challenge for alignment comes from a long history of working together in an externally-led and supply-driven set-up. A recent summary of lessons learned, from the Organization for Economic Cooperation and Development’s Development Assistance Committee (OECD/DAC) Peer Reviews, on supporting partners to develop their capacity, calls on the DAC donors to “take partner ownership and leadership seriously”. The summary concludes that instead of focusing on donors’ domestic accountability and getting the job done with short term results, a change of mentality is necessary for longer term sustainable capacity development results.

Use of external expert inputs (technical assistance) for demand-driven evaluation capacity development

Capacity development implies intentionality to strengthen capacities. For example, technical assistance may strengthen country capacities only if it is clearly oriented towards this aim, by engaging country stakeholders in the evaluation process, creating space for learning by doing (i.e. in joint country-led evaluations), and facilitating access to knowledge. Capacity development-focused technical assistance is therefore demand-driven rather than supply-driven, and it
**Initiatives and institutions for national evaluation capacity development have rapidly increased both in the public sector and in civil society.**

focuses on the process and outcomes rather than the inputs of technical assistance.

Guiding principles to be applied when external expert inputs (technical assistance) for evaluation capacity development are provided include:

- Technical assistance should be accountable to the national institution rather than an external stakeholder.
- The Terms of Reference of externally appointed staff should be determined by the national institution jointly with the provider of external support.
- The technical assistance should be jointly selected and managed by partners.
- Use of national and regional resources should be promoted, e.g. through South-South cooperation.

Figure 5: The shift in paradigm from technical assistance to a capacity development approach

Working jointly for evaluation capacity development led by national institutions

Initiatives and institutions for national evaluation capacity development have rapidly increased both in the public sector and in civil society. There are also signs of growing interest to support these efforts. The expanding number of partnerships for evaluation capacity development brings with it the traditional challenges of coordination and promotion of joint work. It is important that national institutions exercise leadership in coordinating external support. This implies an active role for the national institutions in all decision-making, including in joint management structures with external partners. Coordination is necessary to avoid overlapping activities and to ensure a complementary approach in comprehensively supporting the evaluation system. Clear division of labor between the partners should be the objective. Working jointly with others should always be the preferred option, and carefully analyzed before launching less efficient bilateral activities.

Managing for results

Capacity development monitoring and reporting often suffer if the focus is simply on what has been done and immediately achieved. Reports on training sessions organized and numbers of people trained are familiar to all. A shift to emphasizing medium and longer term results is necessary. The important thing is to report on what actually changes in individuals’ knowledge, skills and will, in institutions and in the operating environment as a result of capacity development efforts. Furthermore, understanding why the selected strategies succeed or fail in causing changes is the input that is needed for future planning. The basis for useful monitoring is in high level planning—clear and measurable objectives accompanied with baseline data. The national institutions’ own results-oriented reporting and results frameworks should be used and developed, or common arrangements made among providers of external support.

Mutual accountability

All partners should be held accountable for their commitments. Consultation and participatory approaches by institutions in planning, implementation, and assessing progress are key strategies for openness and accountability. For external partners, medium-term predictability remains a challenge. Timely, transparent, comprehensive information on external support, including finance flows, is important for national planning processes.

Conclusions

In the past, several stakeholders worked to strengthen National Evaluation Capacities focusing on training individuals, rather than strengthening institutional capacities and an enabling environment. In addition, thanks to the changing external environment and the recognition and involvement of new stakeholders (notably CSOs, including VOPEs) and the need for better partnerships to support nationally-owned and -led evaluation system, the lack of a shared framework for National Evaluation Capacity Development has become evident. This article is a step towards addressing this weakness, as a first step to create better networked partnership, as in the example of EvalPartners.
There is some ambiguity about the key concepts associated with National Evaluation Capacity Development (NECD). To help clarify their usage, the UNEG defines four concepts fundamental to NECD. These comprise:

1. the distinction between ‘evaluation’ and ‘monitoring’;
2. ‘performance management’ and ‘results-based management (RBM)’;
3. ‘national evaluation system’ and ‘monitoring and evaluation (M&E) system’, a term often used in its place;
4. in broad terms, what is implied for a country by ‘national evaluation capacity development’.

Clarifying the terminology of national evaluation capacity development

‘Evaluation’ versus ‘monitoring’—two tools to measure performance

The term ‘evaluation’ is not always well understood. It can take many forms, address a range of issues and potentially serve a variety of uses and users to support a country’s national development goals. It is also often the most cost-effective way to measure ‘outcomes’ associated with projects, programs or policies.

‘Performance monitoring’ can also play an important role in increasing understanding of government performance and measuring progress on national development goals. It requires the identification of relevant indicators of ‘performance’, the establishment of ongoing measurement systems and the availability of credible data to populate the indicators. Once operational, monitoring information can be a ready source of information on ‘outputs’ and (potentially) short-term ‘outcomes’, as well as providing an important input to the conduct of an evaluation.

Monitoring though does not provide the depth of understanding of ‘performance’ that can be gleaned from an evaluation. Also, evaluation is generally needed to measure higher-level results (‘intermediate and longer-term outcomes’ and ‘impacts’ of public sector interventions). But, being a discipline based on social science research methods, evaluation requires a higher skill set and level of experience that is often lacking or in short supply in a country context. Capacity building efforts need to recognize the above distinctions and the importance of strengthening both as key components of an effective and sustainable national system.

‘Performance management’—‘results-based management’

These two terms can be considered interchangeable, both referring to ‘a management strategy focusing on performance and achievement of outputs, outcomes and impacts’. Its broad objective is to advance sound management practice and strengthen accountability and good governance across all organizations.

At a country level, implementing RBM is likely to require some level of public sector reform, which should then be the broad driver for NECD, i.e. introducing a more systematic and results-oriented approach to creating and using performance information in government planning, budgeting, management and decision-making. It means managing programs and organizations with a focus on the...
actual results being achieved. To do this requires: clearly defining expectations about results to be achieved; delivering the programs or services; measuring and evaluating performance; and, making adjustments to improve both efficiency and effectiveness. It also means reporting on performance to citizens. Movement to a RBM system, whether for organizations or countries, is long term and iterative. The two tools of monitoring and evaluation though are critical components.

The terms ‘performance management’ and ‘performance measurement’ are often confused. The foundation of performance management (or RBM) is accurate and timely performance information. Thus, NECD efforts are needed to support development of an information regime that measures, evaluates and reports on the performance of programs; holds managers accountable for achieving results; and ensures unbiased analysis and reporting.

‘National evaluation system’ or ‘monitoring and evaluation (M&E) system’
These are somewhat ambiguous terms and not always well understood. What may be deemed to be a country’s national evaluation system (NES) may in fact have little to do with the practice of ‘evaluation’ per se. In fact, the more widely used term for developing countries is ‘monitoring and evaluation (M&E) system’ and this is often used interchangeably when talking about a country’s NES.

Further, in many instances, the prime focus of the national M&E system is on the ‘M’, with little or no investment in the ‘E’. It is important, therefore, to assess and understand the country context so that any NECD efforts by UN agencies represent the most effective interventions to support the national system. It also underscores the fact that NECD efforts in some cases may need to recognize that there may be little or no operational capacity for evaluation at the outset (though certainly other dimensions to build upon).

BUT, a national evaluation (or, M&E) system needs to be thought of as more than simply the generation of ‘performance’ information. Evaluation (or M&E) is not an ‘end’ in itself. The concept needs to be viewed in terms of a systems approach that recognizes the importance of both an ability to provide sound evidence (the supply side) as well as the capacity within the system for individuals and institutions to use information (the demand side). In supporting NECD, it is thus important to reflect on: (i) who will be carrying out monitoring and evaluation activities; and (ii) who will use the ‘results’ information across the system?

National evaluation capacity development (NECD)
The basis for ‘national evaluation capacity development’ builds on the concepts described above. Underlying it though is the belief that results-based monitoring and evaluation represent public management tools that yield information that can be used to better manage policies, programs and projects; demonstrate progress on national development goals to citizens and stakeholders; and, in the process, to serve as key accountability tools.

NECD needs to be thought of more than simply in technocratic terms though. National ownership implies a particular cultural, social and political context. The driver for establishing a national evaluation (or M&E) system ought to be good governance. And, this could be associated with government modernization initiatives, the proactive introduction of a results-based management framework for the public sector and/or some other form of public sector reform. This could be triggered by political change; or it could be encouraged by the actions and support of international agencies and donors. If it is to be owned by the country, it needs to be linked to the national development plan and integrated into the operations and culture of government institutions and ministries. To be sustainable though, governments must believe in the utility of the NES and understand its benefits. And, to do that, they must eventually own the system.
Perspectives From Around the Globe

AFRICA
- Strengthening National Evaluation Systems
  Samuel Kouakou, Côte d’Ivoire
- The Importance of Strengthening National Evaluation Systems
  F. Kelemework Fasika, Ethiopia
- Developing a National Evaluation System in South Africa
  Ian Goldman
- Political Patterns and Influences in the Supply and Demand of Evaluation and Monitoring in Uganda’s Public Sector over the Past Two Decades, David Rider Smith
- Strengthening National Evaluation Systems
  El Hadji Gueye, Senegal
- Towards an Evaluation Culture in the AfDB: Practices and Challenges, Mohamed Manai, AfDB
- National Evaluation Capacity Development

LATIN AMERICA
- Speaking Truth to Power: The Role of National Evaluation Systems in Democratic Governance, Claudia Maldonado

NORTH AMERICA
- Rebirth, Restoration, Reclamation, and Responsibilities of the Evaluation Function of Africa, Sukai Prom-Jackson, UNDP

Why is it Important to Strengthen National Evaluation Systems?

The Paris High-Level Forum on Aid Effectiveness, hosted by the French government and organized by the Organisation for Economic Co-operation and Development (OECD) in February 2005, and the Third High-Level Forum on Aid Effectiveness (HLF-3), held in Accra, Ghana, September 2–4, 2008, have given a renewed impetus to monitoring and evaluation (M&E).

Many governments in developing countries are devoting considerable effort to strengthening their monitoring and evaluation (M&E) systems and capacities to improve their performance by establishing evidence-based policy making, budget decision making, management, and accountability.

Why Should We Invest Time and Effort in National Evaluation Capacity Development?

It is important to invest time and effort in national evaluation capacity development because capacity development is a strategic investment in the future of the civil service. This can enhance the well-being of citizens and the prosperity of the country. It can also improve government performance.1

We should invest time and effort in national evaluation capacity development in civil society, because nongovernmental actors are thought to play key roles in both the supply of and demand for information and accountability.

The establishment recipient M&E frameworks that function well is also crucial for donors, particularly as they wind down their own parallel M&E systems and become dependent upon recipient systems.²

Many lessons can be drawn from experience in the development of national evaluation capacity. These can be summarized as follows.

**Successes**

- Strong ownership and development-partner support for national evaluation capacity development core reform and sector-based programs, the use of national systems.
- Close collaborative approach between donors and sector ministries, decentralized structures, and with municipalities.
- Effort to further strengthen performance budgeting by governments.
- Significant investment in sector-focused technical cooperation and capacity development initiatives.
- Efforts to engage with civil society.
- Growing number of voluntary organizations for professional evaluation. Indeed, from one in 1981, there are now 135 voluntary organizations for professional evaluation.³

**Failures**

- The use of country systems, however, has remained very low, with tone successful setup.
- In most countries’ national M&E systems, the ‘E’ for evaluation is essentially nonexistent as one of the tools to measure performance. Activities are essentially focused on monitoring.
- M&E efforts are still largely donor-driven.

**Challenges**

The main challenge faced in evaluation capacity development is its unreliable funding and resource base. Most capacity-building organizations depend largely on donor support to conduct their activities. Although a membership fee is charged in some organizations and government supports institutions or ministries, this constitutes a small percentage of their annual budget. These organizations therefore experience periods of very low activity, when no funded project is being implemented.

Most of the voluntary organizations for professional evaluation have weak national anchoring, and they are not called upon by the governments or by the partners in development to carry out evaluations at a national level. For successful capacity development, we must ensure that there is a strong civil society, demanding and advocating for evaluation values and for evidence-based policy making.

Another challenge is the production of international evaluation standards and resources. The norms and standards produced by the network also serve as a form of direct capacity development.

---


that provides partners with resources to build their evaluation knowledge.

It takes time to create or strengthen national evaluation capacities and to build systems for sharing M&E information among relevant ministries.

**Good practices, tips**
Extensive experience with capacity development initiatives has highlighted three major lessons learned:

- Capacity development should be underpinned by the fundamental characteristic of national ownership. Taking a capacity development strategy to scale requires linking the strategy to national and local plans, reform, processes, budgets, and systems.

- Capacity development is about transformations and must address how best to manage change within the existing policy environment.

- A good evaluation capacity development strategy should be tailored to the specific context of each country and institution. It is very important to understand the country context before proceeding.

At the level of communities of practice, we must share evaluation experiences, practices, and information at meetings organized by those who have an interest in engaging in evaluation work.

The strategy for strengthening M&E systems focuses on three aspects:

- Take stock of the situation to identify a scenario for determining growing needs.

- Decide on critical steps to be taken for meeting the above needs, in the form of a plan of action and implementing that plan.

- Define the roles and responsibilities of people at various levels in the organization and build capacity for carrying out those roles effectively.\(^5\)

**Examples of models of national evaluation systems that are performing well**

**a) South Africa case**

In about 2000, M&E in South Africa took on increased prominence.\(^4\) There had been M&E activity since 1995, but it was not coordinated. In 2005, the South African cabinet approved recommendations from the President’s Office on “an implementation plan to develop a monitoring and evaluation (M&E) system.” The system was to include functions such as monitoring; evaluation; early warning; data verification; and data collection, analysis, and reporting.

The South African Constitution places a premium on a responsive, development-oriented public service, effective and accountable stewardship of public resources by the executive government, effective oversight by parliament and the nine provincial legislatures, and public participation in policy and implementation processes.

---


5 “Design of Monitoring & Evaluation System for Strengthening the Sri Lankan Vocational Training System by GTZ SVTP Project—Colombo.”

All of these factors have led the South African government to recognize the need for a government-wide M&E 7 system.

b) Australian case
The election of a reformist labor government in 1983 provided an environment favorable to evaluation. A series of public-sector management reforms were implemented in the first several years of the new government. One aspect of these reforms was the desire to let the managers manage by a devolution of powers and responsibilities that reflected the philosophy that public sector managers would be strongly encouraged to improve their performance if they were provided greater autonomy and the potential to manage their departments with fewer central agency controls and less interference. The cabinet’s agreement to the evaluation strategy 8 was expressed in a formal cabinet decision.

c) Canada case
Evaluation in public sector management extends back to 1969 in Canada, 9 with the initiation of formalized and centralized evaluation practices. The introduction of the practice of evaluation was inspired by the notion of letting the managers manage; that is, allowing deputy ministers of federal government departments to assume greater responsibility for their departments and programs. But they were also to be held accountable for the performance of those programs and the prudent use of public funds. The model is based on a strong central management board that oversees and holds deputies accountable, and one of the mechanisms intended to do this is the performance evaluation.

The 2009 Evaluation Policy has put renewed emphasis on evaluating all direct program expenditures of government and, in the process, focusing on the issues of program relevance and performance—that is, effectiveness, efficiency, and economy. The intent is that evaluation will be

---

better used, by both departments and the central agency, to address higher-level issues, and, ultimately, to better integrate program effectiveness information into decision-making processes, especially expenditure-management decisions.

d) Columbia case

Colombia’s national M&E system (SINERGIA) is one of Latin America’s more visible efforts to institutionalize M&E at the government level.

Established in 1994, following the mandates of the National Constitution, SINERGIA,10 the national system for evaluation of management and results, has endured numerous periods of progress, stagnation, and setbacks because of the institutional, political, and fiscal challenges faced by the country over time. Nonetheless, SINERGIA has attained a significant level of consolidation and customization and is often referred to as a good practice system by multilateral organizations, donor agencies, and academic analysts.

Colombia’s M&E systems qualify as good practice because:

- The quality of their M&E work is generally reliable.
- The monitoring information and evaluation findings they produce receive intensive use.
- Limitations of relying on laws, decrees, and regulations.

Finally, it appears that many factors, such as the strength of leadership for the reform, the resources provided to support it, the establishment of routine rules and procedures, and incentives for the utilization of M&E information, are also crucial to a system’s success.

The new global reality

Over the past two decades, the World Bank and other international financial institutions, and more recently some prominent UN agencies, have devoted resources to building evaluation capacity in selected countries. However, in a world where resources are tighter in all global situations, the context for national evaluation capacity development has changed for UN agencies and partner countries alike. The new global reality is that UN agencies, both on their own and under the coordinating body of the United Nations Evaluation Group, will need to be quite selective in searching for the most cost-effective approach to supporting national evaluation capacity development.

Along the lines of the recent Busan Partnership,11 civil society organizations (CSOs) can and should play a central role in advocating for transparency in the allocation and expenditure of the public budget, accountability for the implementation of public policies, and the strengthening of the demand for the use of evaluation to inform evidence-based policy making, and the capacities of qualified evaluators to produce credible and useful evaluations based on national and international evaluation standards.

In Côte d’Ivoire, many actions and activities are being performed to foster national evaluation capacity development at the national level. These include the following.

11 BUSAN Partnership for Effective Development Cooperation Fourth High-Level Forum on Aid Effectiveness, BUSAN, Republic of Korea, 29 November–1 December 2011.
a) The Ivorian M&E Network (RISE)
RISE has gained official government recognition and is present at all the major meetings organized by the African Evaluation Association (AfrEA).

One of the most recent outcomes is the pre-qualification of RISE for the national association competition organized during the sixth conference of the African Evaluation Association in Accra in January 2012.

Mr. Samuel Kouakou, on behalf of RISE, participated in the 10th Biennial Conference of the European Evaluation Society (EES), which was held in Finland from 1 to 5 October 2012. This activity signified the involvement of RISE in major international meetings on evaluation.

Many interviews have been made available in national newspapers and radio broadcasts advocating for decision makers and the strengthening of evaluator capacities.

Finally, a case study presented by Mr. Samuel Kouakou, on behalf of RISE, has been selected for inclusion in the International EvalPartners Forum, organized on 3-6 December 2012 in Chiang Mai, Thailand. The aim was to strengthen the evaluation capacities of civil society organizations in general, and voluntary organizations for professional evaluation in particular.

The next steps will be:

- Ivorian Evaluation Days will be organized.

b) The national institutions

- Organization of a workshop to strengthen the technical capacity of sectoral ministries and leaders of departments in charge of planning and statistics, operational focal points in the implementation of the National Development Program (NDP 2012–15) and to the methods and foresight tools, strategic planning, programming, budgeting, and M&E (PPBSE cycle);

- The implementation of the M&E system of the the National Development Program (NDP 2012–15) is still in process;

- The design of the M&E system of the National Agricultural Investment Program (PNIA 2012–15) by the Ministry of Agriculture through the Directorate of Evaluation and Projects (DEP) is still in process;

(c) The International level

As a member of the EvalPartners Task Force on the enabling environment, I am deeply involved in the implementation of activities to advocate and promote evaluation and evidence-based policy making at the international, regional, national, and local levels, and in designating 2015 as the International Year of Evaluation (EvalYear).

Many other capacity-building activities will be conducted by the European Evaluation Society (EES) Working Group on Sustainable Development, as well the International Development Evaluation Association (IDEAS).
Working Group on Certification for Development Evaluators, to which I belong.

**Urgent issues in national evaluation capacity development**

**a) The issues of certification of evaluators**

The profession of the evaluator is young. There is no professional order of evaluators. There are national evaluation associations in several countries, generally industrialized countries.

There are many incompetent evaluators, just as there are many seasoned professionals. This makes it difficult for those who set out to hire an evaluator to know who is qualified to do the job and who is not.

The field would benefit from evaluator competencies in four primary ways: (a) improved training, (b) enhanced reflective practice, (c) the advancement of research on evaluation, and (d) the potential for continued professionalization of the field.

The newly accredited CLEAR centers in the Latin America, Africa, and Asia regions that are planning to offer university certifications in the future will assuredly help address these issues.

**b) The issues of enabling environment**

We must ensure that:

- Public administration is committed to transparency and managing for results and accountability through results-based public budgeting and evidence-based policy making.
- There is legislation and/or policies, including public budget, to institutionalize evaluation systems.
- There are national organizations for evaluation professionals.

**Key elements for successful national evaluation capacity development**

In the past, evaluation capacity development mainly focused on strengthening individual capacities. However, it is clear that capacity development should be based on a systemic approach that takes into account three major components: the individual, institutional, and, external enabling environments.¹²

Successful national evaluation capacity development must be based on country-led evaluation systems. Therefore, international organizations and donors should always make sure their own evaluation function strengthens rather than undermines national evaluation systems and capacities.

In the past 15 years, the potential of major national and regional voluntary organizations of professional evaluators grew, as they rose in number from 1 in the 1981 to more than 135 in 2013.¹³

Factors related to the complementary learning strategies that reinforce use of the training include:

- Learning by doing and mentoring: It can be an effective follow-up of the training.
- Self-education: This can be an effective strategy for evaluation capacity building.


• Learning by sharing: This is a powerful way of deepening evaluation knowledge and can be considered one of the best possible outcomes of evaluation training—that is, people using their new learning to teach others.

**Actions that would strengthen evaluation capacity development**

A country strategy for national evaluation capacity development should strengthen the enabling environment by supporting:

• The commitment of public administration to accountability, transparency, and managing for results, including strengthening results-based public budgeting and evidence-based policy-making.

• Legislation and/or policies to institutionalize national evaluation systems should be designed, adopted, and implemented.

• National evaluation organizations should be supported to enable them to foster national demand for and supply of evaluation, including by setting national evaluation standards and norms.

At the institutional level, a strategy for evaluation capacity development should strengthen the corporate evaluation culture through institutional commitment to learning from evidence produced by evaluation systems. It should also support evidence-based decisions and demand for accountability and establish a protective environment that removes repercussions to careers.

Finally, parliaments and civil society organizations can also play a major role in advocating for the use of evidence in policy implementation.

The support of mass media can also make evidence available to citizens, and citizens can demand that policy makers make more use of it.

A good initiative that will strengthen evaluation capacity development is the organization Parliamentarians Forum on Development Evaluation in South Asia. The forum was initiated by a small group of parliamentarians aiming to establish national evaluation policies in the South Asian Association for Regional Cooperation region to ensure aid effectiveness, achievement of results, and sustainability of development.

---

Mr. Samuel Kouakou is Deputy Director of Evaluation and Projects Sustainability in the Directorate of Evaluation and Projects of the Ministry of Agriculture, Côte d’Ivoire.

He has extensive experience in evaluation in public and private service at the national and international levels. He is an active member of the African Evaluation Association, International Development Evaluation Association, European Evaluation Society, and the Ivorian M&E Network (RISE). He holds a diploma in international environmental law from the United Nations Institute for Training and Research (UNITAR), a master’s degree in rural engineering, as well as several post-graduate certificates in the areas of integrated water resources management, gender mainstreaming, and results-based M&E.

Mr. Kouakou is currently e-learning visiting lecturer at the International Institute for Water and Environmental Engineering. He earned his e-learning tutor certification from the Association of Universities of the Francophonie.

---

14 [http://www.mymande.org/evalyear/parliamentarians_forum](http://www.mymande.org/evalyear/parliamentarians_forum)
The Importance of Strengthening National Evaluation Systems - *from Ethiopia*

*F. Kelemework Fasika, European External Action Service-Addis Ababa*

**SINCE THE INITIATION** of Poverty Reduction Strategy Papers (PRSPs), poverty reduction has become the main focus of the national policies and strategies of most developing countries. Moreover, this focus has been shared by regional and global development players for quite some time. Although much has changed over time, poverty and inequality remain challenges, and the need for strong and sustainable growth is still high on the development agenda.

Lack of adequate capacity is considered a factor holding back progress in the attainment of the goals of national poverty reduction strategies, as well as global commitments, including the Millennium Development Goals (MDGs). The need for development effectiveness had been part of the global development discourse. This led to the formulation of the Paris Declaration on Aid Effectiveness in 2005, and the agenda was reviewed and further strengthened through the Accra Agenda for Action and the Bussan Partnership Agreement.

Developing evaluation capacity is a cornerstone in building a holistic capacity to design and implement effective policies and strategies. It provides insight on the current situation, helps identify needed improvements, and generates lessons to enable future progress. Based on this understanding, the objectives of developing evaluation capacity entail promoting efficiency and effectiveness of development interventions; enhancing implementation capacity to respond to the growing needs for service delivery, job creation, and the like; and informing dialogue (both on policy and on evaluation capacity development). Through evaluation, governments can hold their civil services accountable for effective use of development resources, as well as for their performance. Evaluation can also help governments to be more accountable to the citizens and gain legitimacy.

For quite some time, whatever the level of capacity, every developing country had a system to monitor and evaluate its development efforts. However, the focus of such systems had been on inputs and processes rather than on outcomes and impact. Apart from measuring inputs and processes, the effectiveness of any national policy and strategy depends on learning lessons from what works, what doesn’t, and why. Collecting
evidence and measuring results through evaluation are essential in providing insight and new perspectives to inform policies and strategies. Moreover, as the aid effectiveness agenda came into being, progress brought about by development interventions was measured in terms of results and impacts. The role of monitoring and evaluation (M&E) as a tool to inform the design of national plans and strategies, assessing results as well as measuring impact and changes, has increased tremendously. Measuring development effectiveness, learning lessons, and adjusting future action based on evidence generated has become increasingly important. National evaluation systems play a key role in this regard.

Based on the discussion above, the need for strengthening national evaluation systems emanates, first: domestically, from governments with an interest in becoming more efficient in optimizing resources and becoming more effective in achieving development targets, as well as discharging the accountability function to their citizens and, second, from international commitments and principles, including the MDGs; Paris Aid Effectiveness Agenda; Accra Agenda for Action; and, most recently, the Bussan Partnership Agreement.

To become more efficient and effective, as well as to discharge their accountability function to citizens, governments, civil society organizations, the private sector, and other relevant stakeholders need to work together to develop local evaluation capacities in all sectors and at all administrative levels. Evidence gathered through M&E systems primarily helps government to demonstrate the effectiveness of development policies and strategies, mobilize resources, use evaluation results to campaign for election, and obtain legitimacy. Evaluation can also help governments to identify their strengths and weakness, as well as learn lessons to enhance performance. The second reason that stronger evaluation capacity is needed emanates from the interest of development partners. They need to be able to track the effectiveness and efficiency of development interventions and ensure accountability in the use of development resources provided to recipient countries based on project, program, or budget-support interventions. International commitments and principles, including the MDGs, Paris Aid Effectiveness Agenda, Accra Agenda for Action, and Bussan Partnership Agreement, which are endorsed by both donor and recipient countries, are designed to promote development effectiveness and mutual accountability. These global commitments focus on development results and encourage countries that receive development assistance to create transparent results frameworks to monitor progress, evaluate impact, and highlight strategic issues that require policy decisions and become more effective and efficient as they work toward meeting their development goals. Donor countries work to enhance M&E capacity by providing technical know-how and financial support.

Challenges

Delivering on results and achieving faster economic growth and social development across the social and economic sectors—at both the local and national levels—in a decentralized system requires a carefully designed set of policies and strategies with strong and harmonized M&E systems. The effectiveness of any national policy and strategy depends on learning lessons from what works,
what doesn’t, and why. However, answering such questions depends on creating and maintaining sound and harmonized national M&E systems in various administrative levels and sectoral areas, which poses both technical and administrative challenges. These must be addressed based on study and understanding of successful experiences from countries with similar contexts.

The relevance of development policies and strategies fades with time and with changes in the country context. This relevance, as well as efficiency and effectiveness, depends on periodic assessment, review, and amendment. That is why evaluation is vital to our ability to adjust policies as we progress. M&E is, therefore, an essential element in development. Policies and programs have uncertainties. Critical M&E is needed to enable timely adjustment to enhance their relevance, efficiency, effectiveness, impact, and sustainability. Results-guided learning is essential for developing countries such as Ethiopia. To facilitate this, governmental and nongovernmental evaluation services need to work together.

Since the adoption of the Poverty-Reduction Strategy Papers, results frameworks and M&E systems have become routine annexes. It is true that almost all the organizations providing services in Ethiopia have some kind of built-in evaluation as part of their process of planning and/or operations. But this implies two important things. On the one hand, it indicates that evaluation is commonly used. On the other hand, it leads to a general understanding that evaluation is common, and its significance has knowingly or unknowingly been eroded, and in some organizations it has been nothing more than a cliché. Such practices risk overshadowing the real benefits of evaluation.

However, various challenges are hampering the ability to provide the accurate evidence required to fine-tune policies and make evidence-based decisions.

The main challenges are:

- Proper awareness and understanding of the benefits of evaluation: although there is an increasing trend of awareness, and M&E units are becoming common in every institution, the significance of evaluation has been knowingly or unknowingly overlooked, and this overshadows the real benefit.
- Technical know-how on evaluation: a knowledge gap in the technical application of evaluation principles, techniques, and tools, as well as the design of appropriate M&E systems.
- Harmonized M&E processes, guidelines, and structures with details of data collection and analysis methods, data quality assurance, and feedback mechanisms from the local to the national level.
- Strong statistical systems and technologies that deliver on the data needs of M&E systems.
- Coordination and harmonization among evaluation capacity development approaches.
- Existence of different types of M&E systems, which makes data aggregation difficult.
- Indicator development: limited focus on impact; more focus on processes and outputs.
Role of Professional Associations in Evaluation Capacity Development

The Ethiopian Evaluation Association (EEvA) has been established and legally registered in Ethiopia since 2008. Its vision is to become a center of excellence in competency and standards of M&E in Ethiopia and in Africa. Its mission is to engage and enable M&E professionals in the country to contribute to evidence-based decision making and positively influence the socioeconomic development of the country through knowledge generation, awareness creation, capacity building, advocacy, and demonstrating model service in evaluation. Members of the association are M&E professionals working in the government, nongovernmental organizations, academia, and the private sector in the country.

As part of its mission to develop evaluation capacity in the country, the association is currently providing training in M&E to professionals working in the governmental and nongovernmental sectors. The association is also working toward promoting professional, quality evaluation practice in Ethiopia by developing evaluation standards to positively contribute to the development objectives of the country. To achieve this, it creates networks with evaluation professionals, disseminates evaluation information, and partners with like-minded local and global organizations, among other activities.

EEvA considers its role in strengthening national evaluation capacity as including the following tasks:

- Create awareness and sensitization on the benefits of M&E.
- Provide training and facilitate experience-sharing among M&E professionals in government, nongovernmental organizations, and the private sector.
- Support development of M&E systems.
- Disseminate evaluation information to stakeholders through publications, brochures, and electronic media.
- Conduct project, program, and policy evaluation.
- Develop evaluation standards.

The Way Forward

**Leadership & ownership of evaluation capacity development**

The most important aspect of evaluation capacity development is the issue of ownership and leadership to create sustainable capacity. The vision as well as the management of the process should be owned and led by the country; development partners can play a support role. Owners should lead the evaluation capacity development need identification, the design of interventions, and management of the process. Partners, however, can create incentives and resources. Nevertheless, the main incentive to develop evaluation capacity should come from the inside in the form of, for example, the need to improve performance and accountability mechanisms.
Evaluation capacity development can only be sustainable when the weaknesses and gaps in the existing capacity are well understood and the gaps that need to be filled are identified—most importantly, by the owner. This can be contemplated through assessment of existing capacity. Such a starting point provides tremendous room and power for change and can serve as a driver for the process. Moreover, understanding the current situation is a basic precondition to the development of a vision for improved capacity in evaluation.

Having a clear vision and understanding the need for improvement provides the confidence to change and offers a clear road map to navigate the process. When such a vision is present at the political level, it provides enormous political will and the required resources to support the process.

As in most developing countries, it has been observed in Ethiopia that inadequate attention is given to evaluation. In addition, other challenges have been noted, including unclear roles and responsibilities among evaluation units at different levels, lack of commitment, irrelevant and low-quality information, poor accountability for failures, lack of integration of actors in the evaluation system at different levels, problems of mainstreaming lessons drawn from evaluation results, narrowing of the scope of evaluation, poor evaluation capacity, and lack of integration of different components of the system.

Strategies to build evaluation capacity:

- Create political champions for M&E.
- Promote engagement of civil society organizations, academia, and research institutes in evaluation capacity development.
- Link the benefits of M&E to effective policy making, enhancing government efficiency in optimizing resources, and delivering more results.
- Strengthen donor government partnership and dialogue.
- Enhance the implementation of joint donor-government aid effectiveness commitments and implementation mechanisms.

In general, key elements for successful national evaluation capacity development include the following steps:

- Strengthen training and education in evaluation.
- Sensitize and create awareness among policy makers, implementers, and the citizenry.
- Provide technical support for development of M&E systems.
- Enhance donor-government collaboration and dialogue.
- Organize events to create public awareness of evaluation, such as an evaluation week and a learning event to discuss, debate, and share lessons and experience, as well as to identify and address challenges in promoting evaluation.

Actions that would strengthen national evaluation capacity development include:

- Build the capacity of individual evaluators.
- Strengthen evaluation systems and structures.
- Develop harmonized evaluation frameworks, guidelines, and manuals.
Developing a National Evaluation System in South Africa

Dr. Ian Goldman, Head, Evaluation and Research in the Department of Performance Me&E, South African Presidency. ian@po-dpme.gov.za

Background

DURING THE 2000s there was a growing interest in monitoring and evaluation (M&E) in South Africa and the M&E role in the presidency began to strengthen. In 2005 the cabinet approved a plan for the development of a government-wide M&E system. It was envisaged as a system of systems, in which each department would have a functional monitoring system, and the necessary information could be extracted.

In 2007 a policy framework was published to guide the government-wide M&E system. It included the need for frameworks for program performance information, the quality of statistical data, and evaluation, and in the process sought to strengthen the linkages among the presidency, the treasury, and the National Statistics Agency. Policy frameworks were developed for the first two elements in 2007, but not for evaluation.

The government that came to power following the 2009 elections faced a number of challenges, which included 1) persistent poverty and inequality, 2) widespread service delivery protests at the municipal level, and 3) loss of some political support in the elections. These challenges resulted in greater willingness of the ruling party and the government to be frank about the poor quality of public services, corruption, and other governance problems, as well as a political consensus to improve government performance, including a greater focus on M&E. A Ministry of Performance M&E was created in the presidency in 2009, and a Department of Performance M&E (DPME) in January 2010. In addition, an advisory body, the National Planning Commission (NPC), was established in the presidency to focus on a long-term 2030 plan.

1 The first part of the paper draws heavily from Goldman and others 2012.
DPME started by building on the United Kingdom model established by Tony Blair’s administration, focusing on a limited number of priority outcomes, developing ministerial targets for these, plans to deliver them, and then working with relevant departments to unblock delivery of these plans. The next area of work—developed based on the experience of Canada—was the assessment of management performance of departments, the Management Performance Assessment Tool (MPAT). In 2011 the focus shifted to evaluation. Visits were made to Mexico, Colombia, the United States, and Australia. The experiences of Mexico and Colombia were used to develop a National Evaluation Policy Framework (NEPF), which was approved by cabinet in November 2011, and the experience of Australia pre-1996 was used to develop the model for a National Evaluation Plan.

The Approach for the National Evaluation System

Historically, work around evaluation in South Africa was developed in the nongovernmental sector in response to donor requirements (Podems, Jacob, and Goldman 2013). Some departments started developing evaluations in the 2000s, but work was sporadic, with no national system or standards. Hence, although most departments have M&E units, in practice the work that has been undertaken is primarily monitoring. According to the 2013 Management Performance Assessments just released, only 17 percent of departments are undertaking or planning to undertake evaluations. As a result, departments are missing out on the opportunity to improve the effectiveness, efficiency, sustainability, and impact of their policies and programs.

The design of the system as envisaged in the NEPF is seeking to:

- Improve program performance.
- Strengthen accountability,
- Improve policy and decision making.
- Improve knowledge of what works and what doesn’t.

A strategic approach has been taken, focusing on important policies, program, and plans, and the selected evaluations are embedded in a National Evaluation Plan, with 15 evaluations selected each year during 2013/14. The focus has been on utilization, and all evaluations will be made publically available unless confidential, and they must have an improvement plan, which is then monitored. The approach emphasizes learning rather than a punitive approach. The goal is to build evaluation into the culture of departments, not promote resistance and malicious compliance.

Six types of evaluations are envisaged, while specific evaluations may combine these approaches:

- **Diagnostic**—Identifies the root cause of problems and the options that could be considered for addressing them.
- **Design**—A short evaluation of the design of program by M&E units within departments
to ensure designs are robust, ideally before implementation starts.

- **Implementation**—Reflects on the progress of an intervention and how it can be strengthened.

- **Impact**—Identifies the impact and attribution of interventions and how they can be strengthened.

- **Economic**—Looks at the cost-effectiveness or cost-benefit of interventions.

- **Evaluation synthesis**—Draws out lessons across a number of evaluations.

Evaluations are implemented as a partnership between the department(s) concerned and DPME, and DPME partially funds the evaluations (with an average of US$60,000 per evaluation). An Evaluation and Research Unit has been established in DPME to drive the system and provide technical support, backed by a cross-government Evaluation Technical Working Group. The unit had 15 staff as at 1 September 2013, led by a deputy director general. Guidelines are being developed, as well as standards for evaluation and competencies for program managers, M&E staff, and evaluators. Training courses started in September 2012.

### The Current Situation With Evaluations

The first evaluation to pilot the system was on early childhood development. This evaluation was completed in June 2012 and the recommendations in the improvement plan are being implemented. The first National Evaluation Plan was approved by the cabinet in June 2012, and work is now progressing on seven of the eight evaluations included in the plan, with the final reports approved for two of these, one on the reception year of schooling (grade R), and one on a Business Process Outsourcing Incentive Scheme. Fifteen evaluations for 2013/14 are under way or in the procurement phase, and the 15 for 2014/15 have just been selected. These are primarily evaluation of programs. The first policy evaluation is scheduled for 2014/15, along with a design evaluation of a policy.

A similar process is under way at the provincial level, and DPME is currently working with three provinces on provincial evaluation plans. Three departments have now developed provincial evaluation plans that combine internal evaluations with those proposed for the National Evaluation Plan.

A key element in building the system to date has been developing evaluation standards (based on the Organisation for Economic Co-operation and Development Assistance Committee standards), guidelines, and templates to ensure that evaluations are of sufficient quality and credibility. Thirteen guidelines and templates have been produced, and all are available on the DPME website (www.ThePresidency-dpme.gov.za).

In addition to commissioning new evaluations, an audit has been done to identify existing evaluations undertaken between 2006 and 2011. Eighty-three evaluations have been identified, and a quality assessment tool was developed.
based on the evaluation standards. Based on this tool, 71 of the evaluations “passed” the minimum quality standards. The documents for these 71 evaluations will be made available in an evaluation repository, which is available to the public from the DPME website.

**Building Evaluation Capacity**

As quoted earlier, only 17 percent of departments are undertaking or planning to undertake evaluations, and this low participation rate makes experience around evaluation quite limited. There is also limited training about evaluation (Podems, Jacob, and Goldman 2013). Hence the development of the National Evaluation Strategy (NES) is effectively starting with a clean slate, although it has been able to use the experience of a few departments that have been carrying out evaluation, notably the Public Service Commission and the Departments of Social Development and Basic Education.

Clearly there is a major lack of capacity, and while the National Evaluation Plan and provincial and departmental evaluation plans are creating demand, there is an issue of having sufficient supply capacity to be able to deliver these evaluations, and to deliver them with sufficient quality.

A set of evaluation competences has been developed for program managers, M&E staff in government, and evaluators (also available on the DPME website), and these have been used to develop a suite of training courses. The competences cover:

1. **Overarching Considerations**
   1.1 Contextual knowledge and understanding
   1.2 Ethical conduct
   1.3 Interpersonal skills

2. **Evaluation Leadership**

3. **Evaluation Craft**
   3.1 Evaluative discipline and practice
   3.2 Research practice

4. **Implementation of Evaluation**
   4.1 Evaluation planning
   4.2 Managing evaluation
   4.3 Report writing and communication
   4.4 Improvement

Table 1 shows an example of the first two competencies.
Table 1: Extract from South Africa's evaluation competences

<table>
<thead>
<tr>
<th>Domains/ Descriptors</th>
<th>Program Manager</th>
<th>M&amp;E Advisor</th>
<th>Evaluator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Contextual knowledge and understanding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CK1—Sector-specific knowledge: Knowledge of relevant sectors in South Africa (for example, content, institutions, people, and politics) and can appropriately relate the evaluation to the current political, policy, and governance environments and issues</td>
<td>Can explain relevant sector in South Africa (for example, content, institutions, people, and politics) and can appropriately relate the evaluation to the current political, policy, and governance environments and issues</td>
<td>Can learn about relevant sectors in South Africa (for example, content, institutions, people, and politics) and can appropriately relate the evaluation to the current political, policy, and governance environments and issues</td>
<td>Can explain relevant sector in South Africa (for example, content, institutions, people, and politics) and can appropriately relate the evaluation to the current political, policy, and governance environments and issues</td>
</tr>
<tr>
<td>CK4—Intervention knowledge: Understands the specific intervention that is the focus of the evaluation and how and why it was developed and implemented</td>
<td>Can guide others on the specific intervention that is the focus of the evaluation and how and why it was developed and implemented</td>
<td>Can learn about the specific intervention that is the focus of the evaluation and how and why it was developed and implemented (degree of familiarity needed may differ)</td>
<td></td>
</tr>
</tbody>
</table>

The courses are short, around three days in length, timed at the point in the evaluation cycle at which the outputs of the course are relevant, which gives them a strong learning-by-doing focus. The suite includes:

- **Managing Evaluations:** targets program managers (three days) and M&E staff (two x three days), and focuses on developing good evaluation terms of reference. This course was been developed and rolled out for evaluations in the 2012/13 and 2013/14 National Evaluation Plan, as well as for two provinces that have developed provincial evaluation plans.

- **Deepening Evaluations:** targets both program managers and M&E staff (three days). This training focuses on critically interpreting evaluation results. The course has been designed and will be piloted and rolled out in 2013.

- **Methodology:** based on five guidelines that are being finalized on the different types of evaluations, and focuses on helping M&E staff to select the appropriate type of evaluation design. This will be designed and piloted in 2013.
- **Planning Programs and Design Evaluation** (see below)

- **Evidence-Based Policy Making**: focusing on top managers (see below)

One of the areas that have proved to be very weak is the design of programs, which is a limiting factor in conducting evaluations, but also in itself a major limitation on program performance. A guideline was released in July 2013 on Planning Implementation Programs, with a standard log-frame, requirements for diagnosis, risk analysis, and so on. This is likely to have a major effect in improving program design. Arising from this, another guideline will be developed on Design Evaluation, whereby M&E officials will review the design of new programs to see that they are robust and likely to achieve their intended objectives. A course to support this planning process and design evaluation will be developed and piloted in 2013.

To assist with supply of quality evaluations, DPME has created a panel of evaluation service providers, to be used for the evaluations in the National Evaluation Plan, but likely to emerge as a panel across government. The panel has been expanded to 42 organizations, but will need to be enlarged as it takes on this wider role, and the capacity of the organizations enhanced.

### Building Demand for Evaluation

A major limiting factor for the evaluation system is how M&E is viewed. DPME conducted a survey in 2012 to understand this better (Goldman, Ntakumba, and Jacob 2013). In terms of culture-based barriers, more than half of the respondents (54 percent) indicated that problems are not treated as opportunities for learning and improvement. The survey also revealed that senior management often fails to champion M&E (45 percent), M&E is regarded as the job of the M&E unit rather than the responsibility of all managers (44 percent), there is not a strong M&E culture (40 percent), M&E is seen as policing and controlling (39 percent), and M&E units have little influence. These all point to the challenge in using M&E as a strategic function to inform policy and decision making.

In 57 percent of cases, M&E information had limited or no influence on decision making. Nearly half of the respondents (46 percent) regard integration with policy development as either nonexistent or very limited. Just under half of the departments (48 percent) reported that integration of M&E with budgeting is limited. This lack of integration implies a poor environment for the demand for and use of M&E evidence, since it is likely to be viewed as a standalone activity, detached from other key management processes.

This has several implications for the evaluation system:

- There is a push to make it a demand-driven and learning-focused system to overcome the compliance mentality and make it likely that evaluation results will be used to improve performance.
- Senior managers need to buy into the system and see it as important for their own and departmental performance.
• Other sources of demand need to be stimulated, including the cabinet (which approves the National Evaluation Plan and receives reports on evaluations) and parliament, to facilitate their oversight function.

• It is important that evaluations be made public to stimulate demand from the citizenry and to foster use of evaluation results.

With these implications in mind, DPME has decided to conduct training for senior managers. A case-based three-day course is being developed for top managers (directors general/permanent secretaries and deputy directors general/deputy secretaries), which is being piloted in November 2013. To make it more interesting to the participants, the focus is on evidence-based policy making and implementation, and international experience is being brought in with the support from the Center for Learning on Evaluation and Results (CLEAR).

In addition, intensive work is being undertaken with parliament. DPME actually reports to the Standing Committee on Appropriations. Workshops have been run for the committee, and study tours undertaken—to the United States and Canada in 2012, and Kenya and Uganda in 2013—to build the understanding of the committee around M&E. This has been very successful, and DPME has now done presentations to the chairs of all portfolio committees, to parliamentary researchers, and to specific portfolio committees, to stimulate their awareness around evaluations and how they can use the results to facilitate their oversight role. In June, the first cases were presented to portfolio committees, and this led to requests to departments to submit specific evaluations. DPME is now planning an extensive program of capacity development for parliament to assist committees in using DPME’s M&E information.

The Road Ahead

This paper has outlined where South Africa is in terms of the National Evaluation Strategy. Key issues that will emerge in the next few months include how departments act when evaluation results are challenging, and the implications of this for the evaluation system. The next year will be an interesting one, as 23 evaluations conclude.

Capacity in a number of departments will increase, because they are involved in several short courses and will complete a number of evaluations (for example, rural development, trade and industry, and human settlements). This experience will create a cadre of experienced evaluation managers.

There is an emerging challenge in being able to supply sufficient numbers of high-quality evaluators. DPME conducted training around the South African M&E Association Conference in September 2013, specifically targeting experienced evaluators.

Some of the issues that emerged include:

• Among evaluators from companies, there are areas of weakness in basic research competences, such as using literature reviews to generate analytical frameworks to guide the evaluation.

• Among universities, there are weaknesses in being able to write for a policy audience.
and to generate recommendations that are closely linked to implementation realities.

- The need to write short reports is a challenge. The National Evaluation Strategy requires that all final evaluation reports use a 1/3/25 format—one-page policy summaries, 3-4 page executive summaries, and 25-page main reports—to ensure readability for busy managers and policy makers.

- Bearing in mind the South African context, there is a shortage of experienced black evaluators. As well as being a weakness in itself, this leads to a lack of credibility of some evaluators, with a wariness among policy makers about using (the same) old white male researchers and evaluators.

All of this points to the need to influence tertiary courses in M&E and to take note of the competences needed around evaluation, and DPME is in discussions with tertiary institutions to this end.

References


Dr. Ian Goldman started his career as an agronomist in Mexico in 1978-79, and has since worked in 18 countries in Africa, Europe, and Latin America, and with national, provincial, and local government, as well as in the NGO and small and medium-size enterprise sectors. In October 2009, Dr. Ian Goldman became Team Leader of the Monitoring and Learning Facility of the Programme to Support Pro-Poor Policy Development (PSPPD) and in January 2011 he was appointed a deputy director general in the Department of Performance M&E (DPME) in the presidency, where he is responsible for evaluation and research. In this capacity he is leading the establishment of the National Evaluation System. Many elements of the system have been established, including a National Evaluation Policy Framework, guidelines and templates, evaluation standards, evaluation competencies, and 5 courses; 20 evaluations are under way.
Patterns and Influences in the Supply and Demand of Evaluation and Monitoring in Uganda’s Public Sector over the Past Two Decades

David Rider Smith, UK Department for International Development’s Evaluation Adviser in Uganda and Acting Head of the Capacity and Quality Group in its Evaluation Department

Context

Over the past two decades, considerable efforts have been made to establish a strong and robust basis for assessing public spending and its effects on the development of Uganda and its citizens.

To better understand the linkages, patterns and constraints to growth and change, substantial resources have been spent on establishing good time-series and qualitative data on key socio-economic indicators; public accounts; regular monitoring of public policy interventions; and, on policy relevant research, analysis and evaluation.

In order to understand where this investment in monitoring, evaluative research and evaluation has had the biggest impact on public policy and accountability, it is necessary to examine the relationships between policy, institutions and individuals in the public sphere. Evidence suggests that only when the environment in each of these spheres is conducive and interfaced has the linkage between assessment and policy change been productive.

---

1 David Rider Smith is an Evaluation Adviser with the UK Department for International Development. All views herein are solely those of the author and do not represent the views of his employer.

2 In this paper, evaluative research conducted in the public sector, i.e. led by public or quasi-public sector institutions, refers to analysis not only of trends, but also of causes and potential policy responses. These include the Poverty Status Reports (PSRs) and related analytical products.

3 This paper builds on one presented at the 2012 American Evaluation Association Conference entitled ‘political economy of evaluation in Uganda’s public sector’ and the article by Rider Smith, Nuwamanya and Nabbumba Navengo, 2010. Policies, Institutions and Personalities: Lessons for Uganda’s experience in Monitoring and Evaluation in From Policies to Results: Developing capacities for country monitoring and evaluation systems, UNICEF.
**Normative Framework**

Relevant literature points to the critical prerequisite of a variety of demand-side elements in the use of evaluation and evaluative research in public policy and programmatic decision making. These elements can be categorized into four that pertain to the openness of the political system to evidence and argument; organizational systems that have elements of performance measurement and analysis embedded within; individual leadership where relevant policy makers have an interest in analytical work and/or knowledge of relevant subject matter; and individual evaluation factors relating to the timing and focus of evaluations. While there is no empirical evidence on the depth or extent to which these factors need to be present or active, nor in terms of the combination of which factors may prove to combine most effectively with others to create demand, elements of each are considered to be necessary for effective uptake.

On the supply side, critical factors can also be categorized into four, including a framework of legal or administrative policy for evaluation or public policy research; the presence of systems for designing, commissioning and/or conducting and disseminating evaluations or research; the capacity to evaluation through a strong social science academia and consultancy sector; and in the nascent phases of economic development, the presence of external assistance to finance such analytical work.

Supply and demand elements are not mutually exclusive. While demand is considered critical to uptake, it also relies upon, and can be strengthened by an adequate framework and system for supply.

**Poverty analysis and its impact on public policy**

The Government of Uganda started to produce poverty monitoring data in 1992, through the Uganda National Household Survey (UNHS) reports, and has since updated this information every two to three years. This data, however, did not play a vital role in assessing public policy until the launch of the Poverty Eradication Action Plan (PEAP) in 1997. Heavy investments in education and health service delivery through the PEAP made it necessary for the Government to assess closely the usefulness of these interventions in terms of transforming society welfare. Hence, the policy environment and actors were open to the production of reports that would, in effect, illustrate how far the national policy was being effective.

In 1999, the Government of Uganda designed a poverty monitoring strategy that guided the production of biannual poverty status reports, and associated monitoring reports and publications. The Poverty Status Reports (PSRs) were high quality analytical pieces which drew upon quantitative and qualitative household and facilities survey data to determine the patterns and changes in rural and urban poverty.

---


5 The full normative framework can be found at the paper presented at the 2012 American Evaluation Association of the same name.
The institutions responsible for generating, analyzing and reporting poverty data were critical to this process, not only their effectiveness as stand-alone institutions, but also for the inter-relationships between these agencies. The PSRs were reliant on good, regular statistical data production, and the work of the Uganda Bureau of Statistics (UBOS) was critical in this. UBOS, with substantial financing from the World Bank, initiated a process of producing regular nationwide household surveys on household income and poverty, health status, population trends, and then later, on other economic and public policy issues.

Within the Ministry of Finance, Planning and Economic Development (MFPED), a Poverty Monitoring and Analysis Unit (PMAU) was established, to monitor, analyse and report on data generated on poverty and related issues, including the preparation of PSRs (with financing from UK). This work was supported by the Government-sponsored Economic Policy Research Centre (EPRC). The Uganda Participatory Poverty Assessment Process (UPPAP) was also established in MFPED to provide qualitative data on key social economic indicators and the impacts of selected Government policies on the citizenry. The qualitative data was an integral part of the information used to prepare PSRs.

Each of these establishments: UBOS, PMAU, UPPAP and EPRC had highly qualified, committed and motivated individuals in key positions. The ability to produce high quality monitoring reports in a timely manner for political and administrative consumption reflected not only their individual abilities, but their willingness to work together to deliver demand driven monitoring reports. Within MFPED, a strong working relationship between PMAU and the top management of the Ministry meant that findings from these products made their way into policy and allocative decisions, and in turn, attracted increased official development assistance. Many of the issues raised in the PSRs influenced decision-making at both at the Cabinet and the Parliamentary level, and helped in focusing expenditures in areas that were most meaningful for poverty reduction.

The window of opportunity and practice in the production of PSRs in Uganda reflected the priority and progress made in poverty reduction from the President down, and the relationships and abilities of the institutions and individuals involved.
produced in the second UPPAP carried out in 2002, finding that the PEAP did indeed place greater emphasis on cross-cutting issues, such as environment, and recognized the heterogeneity of the poor in the greater emphasis placed on the decentralized delivery of services, and on district level plans. In a more subtle manner, while the 1998/99 UPPAP raised concern over the negative impact on the poor of cost sharing in health services, this was not addressed in the PEAP of 2000, but was announced by the Government in 2001 during the election campaign (OPM, 2008). More broadly, the timing of the PEAP cycles, revised in 2000 and 2004, did not match with the election cycles of 2001 and 2006, which has resulted in some analytically driven analysis appearing in the PEAP or emanating during or following an election, such as the Strategic Export Initiative, which was initiated in 2001 following the election, but was not evident in the PEAP.

The decline in the PEAP’s influence during its third phase (2004-07) occurred as the public policy debate on development within the country and amongst international stakeholders shifted towards economic growth and enhancing the accountability of the state, in the face of evidence of malpractice and corruption in the use of the state’s resources.

**Shift to budget and performance monitoring for accountability**

Whilst the evidence from household surveys and PSRs began to reveal that the overall poverty headcount was reducing in the country, it was also being noted that growth and development was becoming increasingly imbalanced (MFPED, 2005; UBOS, 2006). As the public purse expanded, based both on a strong and stable economic growth rate, relatively low inflation, and a considerable rise in official development assistance, so too did concerns on the application and accountability for public spending. Efforts to strengthen public financial management included the recognized need to understand how public resources were being applied centrally, and locally under the decentralized system of government and public programme implementation, introduced in 1997.

The emphasis on monitoring shifted away from periodic analysis of poverty trends and causes, and towards the monitoring of budget spending. During the latter half of the last decade, the MFPED introduced a series of reforms to enable Ministries, Departments and Agencies (MDAs) and Local Governments (LGs) to plan and budget annually according to clear budget lines, and against the provision of products and services. Systems have been introduced requiring all MDAs and LGs to report quarterly on spending and progress towards stated output (product and service) targets, as the basis for future financial releases.

This massive reorganization and growth in the administration of Government placed increasing attention on the generation and use of administrative data and statistics, and on the monitoring and oversight mechanisms in place to capture and report on performance information.

---


The political interest and pressure to monitor spending and results has increased since the re-introduction of multi-party politics in Uganda in 2006, and the growing attention of the domestic media and international community to unequal growth, and the incidence of corruption in the use of public resources. The President and other senior policy makers have taken cognizance of these issues, and have placed increasing demands on the public service to improve its stewardship of resources and ensure effective development. The impact of this on public institutions is still unfolding. Efforts to improve oversight in key service delivery institutions (through regular implementation and budget monitoring), and through inspection of service delivery, have increased, though not in a uniform or consistent manner. The former PMAU in the MFPED has been transformed into a Budget Monitoring and Accountability Unit (BMAU) to track expenditure and outputs against budgets and planned activities in a sample of frontline institutions, through direct field monitoring and reporting. Monitoring is focused on the outputs that are agreed and signed up in the Performance Contracts between the Permanent Secretary of MFPED and the implementing agencies. Efforts to reorganize the overall inspection function of Government are underway. A Presidential directive to initiate public fora to hold local Government to account (so called public ‘barazas’) has been initiated by the Office of the Prime Minister, and the UBOS is seeking to expand its mandate to improve the quality of administrative statistics.

Simultaneously, the Office of the Prime Minister (OPM) has reinvigorated its own constitutional role on coordinating the implementation of public policies and programmes through establishing a robust monitoring coordination and oversight function. Building on an early attempt at producing an outcome-based review of the PEAP in 2007, bi-annual Government Performance Reports were initiated in the Financial Year 2008/09 and two-day retreats of all Ministers, Permanent Secretaries established under the President to discuss the performance report, holder portfolio Ministers to account, and propose corrective measures. Since 2011, these bi-annual retreats have been expanded to include all Local Government Council Chairpersons and Chief Administrative Officers. This has expanded the basis of debate around public service delivery.

In this new environment, accountability has become the central concern, with the consequent de-emphasis on generating information for the purposes of understanding causes and generating policy lessons. The considerable differences in practices across Government in the monitoring and inspection of public investments reflects the balance of priorities, incentives and capacities across the sectors, as influenced in part by the international community who continue to invest in certain sectors over others (notably the front line services such as health, education, water and increasingly roads).
The effectiveness of the increased monitoring of public spending has yet to be born out. The regular public presentation of information on the performance of Government has not yet appeared to have impacted on public policy, but has resulted in the greater public focus on the need to enforce accountabilities, and significantly has also revealed the widespread misuse of funds. However, given the breadth and depth of evidence on challenges to public service delivery, the political class and legislative arm have still to make best use of this information in shifting policy directions, reallocating resources to more efficient areas, or in cases of misuse of resources, holding culprits to account.

**Efforts to strengthen the analytical and the evaluative**

The concerted efforts to strengthen monitoring have come at a cost. The practice of public sector evaluation has never been institutionalized in the country, but was reasonably well serviced in the late 1990s and early 2000s through the PSRs and other analytical tools and products. This has led to a deficit in the analysis of results and constraints, and in the identification of policy lessons and choices. Monitoring itself does not answer these questions or address these issues.

Between 2005 and 2008, a review of evaluation in Uganda found eighty-five evaluations commissioned, of which ten were commissioned or co-managed by the Government (OPM, 2009). When reviewing these ten reports in detail, several were found not to meet basic quality standards for evaluation, and were subsequently reclassified as performance assessments or reviews. In terms of coverage of public investments, Government estimated in 2009 that less than 10 percent of projects over the period 2005-08 were being subjected to evaluation. In a sample of Ministries, Departments and Agencies, the same review found little explicit demand for evaluations, aligned with weak organizational capacity and limited use of those that were conducted. In an apparent contradiction, it was also found that there was a perceived need for 'more evaluation' in principal (ibid), reflecting not weak demand per se, but the lack of a clear policy, incentives and resources to evaluate.

Of the evaluations that were conducted during that period, there is little evidence of their impact due to a lack of appraisal by Government or the international partners on dissemination or use. An exception were the evaluations covering the agriculture sector, starting with one covering the Plan for the Modernization of Agriculture (PMA) in 2005, a second looking at the performance of the National Agricultural Advisory and Development Service (NAADS) in 2008, and a third being an impact evaluation also on NAADS in 2009. Each of these independent evaluations gave a broadly positive assessment of progress, with the impact evaluation on 2009 showing positive results on adoption of improved technologies, productivity and per capita incomes. The study (Benin, 2009) also found that between 2004-08, NAADS was associated

---


9 Including donor-financed projects implemented through the public sector.

with an average of 24-53 percent increase in per capita agricultural income of the programme’s direct participants compared to their non-participant counterparts. However, as noted by other commentators, that despite reported successes of NAADS, overall indicators for agricultural growth were not improving (Kjaer and Joughin, 2011). This has presented a problem for Government, and resulted in changing strategies on agriculture, and NAADS, including the renationalization of extension workers, despite the finding from the 2008 evaluation that ‘a return to using public sector extension workers for the majority of services was not a viable option’ (Ibid, 2011). This raises two issues, one relates to the unit of analysis of the evaluation, where NAADS as an initiative may be seen to be relatively successful, but does not take into the broader context, which may be less positive, and hence points at more fundamental structural challenges. The second, reflects the use of evaluative evidence in cases whether the majority of the population, including policy makers, have direct personal knowledge as landowners and farmers, and where the majority of the electorate live of the land, and thus require evidence of efforts to improve their lot.

The paucity of rigorous public sector evaluations has been identified as a constraint to improving the culture of debating empirical evidence in public policy. To address this, the Office of the Prime Minister began an initiative to strengthen the framework and production of rigorous evaluations across the public service. Starting in 2008, OPM led on the design, implementation and dissemination of evaluations of the successes and failures of the PEAP over the decade of implementation, and two (2008 and 2011) on implementation of the Paris Declaration on Aid Effectiveness in Uganda.

The PEAP evaluation process was important in that it was managed by a steering committee composed of representatives from OPM, MFPED and NPA, as well as from the funding donor agencies, and was a good example of how inter-ministry coordination can work, if there was a specific focus or task. It was also important in that OPM understood how the evaluation results could and should be disseminated and acted upon, managing a series of workshops for various stakeholders, writing a white paper for Cabinet, based on the evaluation results and government response. While somewhat supply driven in origin, the evaluation did find an audience amongst policy makers, with the findings and recommendations discussed twice by Cabinet, and have in turn influenced the shape of the PEAP’s successor, the five-year National Development Plan.

Subsequent efforts to strengthen evaluation practice include the development of a national policy on monitoring and evaluation which defines

---

the roles, requirements and practices to embed into the public service, approved by Cabinet in May 2013. The Policy outlines the delineated functions of monitoring and evaluation, and provides a prescription for the establishment of a Government Evaluation Facility (GEF). Preparation for the GEF began in parallel to the Policy in 2010, with the Facility launched in 2011. The Secretariat of the GEF is located at the OPM, and includes the components of a rolling national evaluation agenda determined by Cabinet; a virtual fund to provide reliable financing for evaluations selected; and a national sub-committee on evaluation composed of representatives of Government, academia, NGO sector and the donor community to oversee the design, production, quality assurance and dissemination of the studies.

As of July 2013, the GEF has completed two major public policy evaluations on the effectiveness of Government’s response to absenteeism in the public service and the effectiveness of public procurement and disposals authority and has a pipeline of six further major public policy evaluations, covering a variety of public service delivery related topics, including; the effectiveness of Government’s employment strategy, a comparative evaluation of public and private service delivery; the impact of the land act amendments on illegal land evictions, the impact of aspects of the Northern Uganda Social Action Fund,. Each evaluation is managed by a central coordinating Ministry, either OPM, MFPED or NPA, with evaluations conducted in house, or outsourced to research or consultancy institutions depending on the topic and capacity. All evaluations are subjected to independent reference groups for quality assurance, and Cabinet papers are written containing the findings to facilitate uptake. Government responses are required to all evaluation, building on the experience of drafting a Cabinet White Paper to the independent evaluation of the PEAP in 2008, where the actions from the evaluation were rigorously followed-up.

The strengthening of supply is linked back to demand by senior civil servants and politicians to revitalized some of the coordination structures within Government. A change in Minister and Permanent Secretary in OPM in 2009 led to the renewal of the national coordination framework of committees established by Cabinet in 2003, but left dormant in the intervening period. A three tier structure of committees link Cabinet with cross-sectoral technocrats, and provide a conduit for feeding Government-wide directives down to implementers, and feeding evidence from analysis back. A national monitoring and evaluation technical working group meets bi-monthly and wide representation from across Government, NGO and donor communities. Sub-committees on evaluation, and on monitoring and oversight take up much of the work which feeds back into the working group and on to Cabinet.

This strengthening of Cabinet lead coordination system is ongoing. A feature of Uganda’s public sector governance arrangements has historically been the power of the Presidency and the relative weakness of the Cabinet system. A practical consequence is that Ministers are not subjected to a collective government discipline. There is no clear means therefore of holding MDAs responsible and accountable for their performance. The main lever for collective discipline is the withholding of funding by MFPED, but it often cannot be applied to
A feature of Uganda’s public sector governance arrangements has historically been the power of the Presidency and the relative weakness of the Cabinet system.

core government services, and is ineffective against MDAs with powerful political backing. The moves by Ministers and subsequently the appointment of a new and politically powerful Prime Minister following the national elections of 2011 have been important moves by the Government to fill out the role provided for the Prime Minister in the Constitution. The Committee structure is a major step forward in strengthening the Cabinet’s role in maintaining a strategic and collective demand for performance to which portfolio Ministers and their MDAs are subordinate. It thus provides the space for empirical evidence to be considered and discussed at a high level, and thus give a greater chance for uptake in public policy and implementation.

The withdrawal of donor financing and the increasing role of politics in civil administration

These reforms took a backward step in late 2012 when a case of grand corruption was identified in the Office of the Prime Minister and other parts of Government, resulting in a large-scale withdrawal of donor financial aid and budget support. This had two immediate impacts on the evaluation agenda. First, the majority share of the recurrent and development budget of the monitoring and evaluation activities was financed by donors, and thus the aid freeze had an immediate and big impact on staffing with numbers reducing by over half in the department and the majority of activities stopped, including several evaluations. Second, the credibility of the Office came into question with investigations into senior and mid-level staff. During this period of investigation, court cases and internal reorganization, the instability and lack of finances resulted in the stalling of the evaluation agenda.

Wider questions regarding the legitimacy of the OPM to effectively plays its constitutionally mandated role of leading government business and coordinating government policies and programmes have been raised—though to a greater extent by donors than internally within the public service. With the movement of senior civil servants, including the Permanent Secretary, and progress made Government-wide in addressing a number of financial controls and accountability issues, support is now returning. This includes to the evaluation and evidence agenda. However, the form of this assistance is likely to change. Financial aid is likely to be provided in exceptional circumstances where controls are strong and alternative modes of delivery inappropriate, so donors are likely to return to project-type support through self-managed procurement of technical assistance and services for evaluation. This is likely to achieve the necessary reduction.

12 Contract staff represented approximately 80% of all staff in the Department of M&E, OPM. Of these, 70% had contracts with OPM, all of which were financed through donor funding. With the freeze of official aid to OPM, the Government cancelled the contracts of these staff after two months in December 2012.
in fiduciary risk, but also reduce public sector ownership, and possibly commitment.

As a break point, the corruption case and the freezing of aid enabled observers to see the extent to which the Government is committed to the monitoring and evaluation agenda, having to finance all operations from its own core budget. Government monitoring, the performance report and Cabinet retreats have continued, albeit with some challenges to quality posed by the substantial reduction in trained staff. The evaluation agenda has suffered, with dwindling numbers attending the cross-Government evaluation committee and despite a healthy pipeline of evaluations, slow progress being registered.

The introduction of a new Permanent Secretary to the Office of the Prime Minister in June 2013 will change the landscape again and opportunities for evaluation in the public sector. Early signs of promise are that she has the political support to shake-up the personnel and systems, with a number of changes in the leadership, audit and procurement staff. It will take time to establish herself and tackle a series of issues before turning to efforts to strengthen the evaluation function.

The political space for evidence-informed debate in Uganda appears to be reducing. The open contestation of ideas within the ruling party, particularly amongst the younger members, have been controlled, and new legislation, such as the public order management bill, is reflective of attempts by the Government to control opposition and public rallies. Reforms within the public service remain piecemeal in sequencing and financing, and consequently there are major challenges in education, health and infrastructure provision. In this context, decision-making is increasing based on allegiance or defensive bases, and not to harness innovation or new approaches. This positions evaluation firmly on supply-side within the public sector, with limited opportunities for growth. However, greater demand may exist amongst non-state actors to amplify the voices of the public.

Conclusion

The experience of Uganda over the past two decades has illustrated that the establishment and effectiveness of monitoring and evaluation practices are most acute when policies, institutions and individual actors come together. The composition and balance of these factors shift over time. The poverty-focused analytical work was pre-eminent in the late 1990s/early 2000s at a time when there was political consensus over the required direction of change within the new Government and with strong support from the donor community. Thus, the analytical work had a willing audience within the political class and amongst the country’s supporters.

As the PEAP began to lose traction, and the shift in Government policy moved towards economic growth, so too there was a reduction in the supply and demand of poverty analysis, and a tactical shift with MFPED towards monitoring and expenditure tracking. The number of agencies involved in monitoring and oversight has proliferated, as demand pressures and supply opportunities within the public sector have increased.
In the last three years, notwithstanding the continued focus on monitoring for accountability, the OPM has led the drive towards reintroducing more analytical work in the policy debate through the establishment of the Government Evaluation Facility. Shifts in the political economy of the Cabinet, with a new and powerful Prime Minister and a reinvigorated coordination mechanism appeared to providing a basis and structure through which demand for evidence could be elicited to inform public service delivery.

However, the demand-side conditions considered prerequisites for effective evaluation uptake and use seem to be decreasing more recently in Uganda. The openness of the political space for debate is reducing, as the move towards the next election campaign in advance of 2016 nears. The organizational systems and the individual champions are in some cases still present, but are unable thrive or function as effectively as they should when undermined by corruption cases, potential loss of legitimacy and in the context of continuingly poorly performing public services. The opportunities for evaluation to influence decisions may now lie primarily outside the public sector and with the public themselves, supporting citizens to demand better services and rights.

David Rider Smith is the UK Department for International Development’s Evaluation Adviser in Uganda and Acting Head of the Capacity and Quality Group in its Evaluation Department. Prior to joining DFID in 2012, he worked for five years in the Uganda Bureau of Statistics and the Office of the Prime Minister in Uganda as an advisor to the Government, leading on the consolidation of performance monitoring, Cabinet reporting, strengthening district statistics and setting up the Government Evaluation Facility. He has previous worked as an evaluation specialist in the Evaluation Office of UNDP, as a researcher at the Natural Resources Institute, and as a survey manager for USAID-supported programmes in Mozambique. His background is in geography, sociology and development economics.
Strengthening National Evaluation Systems—
from CLEAR Francophone Africa

Dr. El Hadji Gueye, BA, MA, PhD, Coordinator of CLEAR Initiative for Francophone Africa

STRENGTHENING NATIONAL EVALUATION systems is important for good governance, improvement of the effectiveness of public policy, and better management of development work. Transparency in the allocation and management of state funds and official development assistance, the obligation to be accountable, and the emergence of a culture of results are some of the many reasons that justify the importance of creating and encouraging an evaluative culture in the countries of Francophone Africa.

The demand is there. If it is not supported by national governments, it is strongly encouraged by technical and financial partners. It is increasingly a strong conditionality in official development assistance. However, the evaluation of public policies, programs, and projects to draw lessons that can help guide development actions is not an easy exercise.

Learning from the implementation of development activities is not systematic. Ineffective approaches and methodologies and lax procedures are repeated through ignorance or lack of communication between development actors. Countries suffer from the lack of national experts capable of undertaking conclusive evaluation studies. It is therefore crucial to develop a comprehensive program of capacity building or empowerment of development actors for better appropriation of the culture of evaluation and harmonization of approaches, concepts, and methodology. This can improve the leadership of the states, and inevitably requires the establishment of a nationwide results-based assessment system.

It is important to invest time and effort in NECD because technical capacity to provide good-quality assessment is inadequate. However, in a context marked by weak political commitment to a national system of monitoring and evaluation (M&E), poor-quality assessments discourage government initiative. It is therefore urgent that we create a critical mass of evaluative competencies to root evaluation in everyday practice and
Unfortunately, little has been implemented successfully. This is often due to a lack of political will. To achieve ownership of evaluation as a real tool that helps in decision making, such behavior should be encouraged and supported by the central government, but also by the parliament, which has a fundamental role to play.

Lessons learned with respect to NECD

Successes, failures

Unfortunately, little has been implemented successfully. This is often due to a lack of political will. To achieve ownership of evaluation as a real tool that helps in decision making, such behavior should be encouraged and supported by the central government, but also by the parliament, which has a fundamental role to play.

The establishment of an enabling legal framework is a prerequisite. But on its own, it is not enough. It should be followed by raising the awareness of development actors at all levels by training evaluation professionals and creating a motivating institutional framework for more rigorous evaluations.

Another source of failure lies in the concept of the independence of the evaluator. The unbridled power struggles (with strong political involvement) that take place force some evaluators to make unethical compromises to please politicians, donors, and civil society leaders. National assessment bodies must have the autonomy necessary for the exercise of their duties.

Challenges

The main challenge is successful advocacy for a good evaluation culture in all Francophone African countries. We need to bring politicians and holders of public power into the evaluation dynamic; make evaluation an integral part of the policy-making, planning, and budget cycle; and encourage the use of assessment results and good governance.

Another challenge is how to successfully deal with conflicts of interest and the influence of
political power and not allow the results of the assessment to be diverted by the strong involvement of politics.

Good practices, tips

- Implement a favorable and motivating institutional framework for evaluation at all levels of public administration.

- Develop national assessment policies and monitor their rigorous enforcement.

- Raise the awareness of managers in public administration about the culture of evaluation.

- Support institutions specialized in training, consultancy, and capacity building in evaluation.

- One should focus on action, practice, and field work rather than use the traditional approach.

- Take baby steps and consolidate gains before moving to the next level.

- Involve civil society and ensure that there is a mechanism for citizen participation.

- Create international networks of evaluation professionals to promote the sharing of experiences and exchanges with a view to harmonizing practices.

Examples of models of national evaluation systems that are performing well

Ghana’s example has a lot to teach us, as does the South African model. Benin’s model can also be used as an example, although much remains to be done there.

Models for consideration:

- The Public Expenditure Management and Financial Accountability Review (PEMFAR), based on the public expenditure and financial accountability (PEFA) indicators

- SNISE (National Integrated Monitoring and Evaluation System), which is the framework for all government planning.

Institutions involved in NECD include the following:

- **Benin**: Institute for Empirical Research in Political Economy.

- **Benin**: The Office of Public Evaluation (BEPP) in the Prime Ministers Office, the Department for Projects and Programs Monitoring in the Ministry for Development.

- **Burkina Faso**: The Department of Training and Educational Innovation in 2IE, which provides continuous professional development programs in M&E for professionals and organizations.

- **Burkina Faso**: International Society of Sustainability Professionals.

- **Côte d’Ivoire**: Global Partnership for Education.

- **Gabon**: ISTA.

- **Senegal**: African Centre for Advanced Studies in Management (CESAG), which provides training and seminars in M&E of public policies and programs.
• **Senegal**: The CLEAR Initiative for Francophone Africa; CESAG, which provides training for capacity building in development of results-oriented M&E, support and guidance to states on the development and implementation of policies in M&E at national and sectoral levels, innovations and case studies.

**The road ahead**

**Urgent issues in national evaluation capacity development include the following:**
The independence of national evaluators and the positioning of the government structures in charge of evaluation.

We should:

- Work for the definition of a national assessment policy for all Francophone African countries.

- Establish a coordination unit for all evaluation activities.

- Create a regional crucible for meetings and exchanges of all professionals to share best practices.

- Create planning and M&E units in the central administration and in ministry departments.

- Extend advocacy for an evaluation culture to civil society and local elected officials.

- Make provisions in the state budget for the financing of capacity building in evaluation.

- Develop an effective process for selecting the best professional profiles, active in the results-production chain and in the conduct of policies, projects, and programs and equip them with the approaches and methods needed to conduct successful evaluation missions.

**Key elements for successful national evaluation capacity development**

First, commitment by the state to efficient national M&E and effective involvement of the state, alongside donors. This should be followed by the definition of a legal framework for the systematic evaluation of public policies, projects, and programs.

Other useful steps:

- Demystify evaluation and institutionalize it.

- Create national structures for evaluation at the country level or revive existing structures; associate them with the actions of CLEAR.

- Establish structures to promote evaluation and organize annual symposia, highly scientific lectures, and discussions on methods and concepts for the best use of the tools.

- Contribute to the development of national policies based on lessons learned.
• Introduce evaluation in academic training programs.

• Create a dynamic regional network of evaluation practitioners.

**Actions that would strengthen evaluation capacity development**
The fundamental measures are:

i) Advocate for a culture of evaluation among governments, policy makers, and development actors at all levels.

ii) Carry out training and capacity building to create a panel in each country consisting of local experts, able to adequately conduct assessment missions.

iii) Provide support and advice to the bodies involved in the conduct of development policy and programs, as well as those in policy-making positions in the state machinery and in parliament, civil society, and decentralized agencies.

iv) Create a degree course in evaluation and provide funding for applied research in the field of evaluation in order to change practices; develop new tools that are more adapted to the particular context of Francophone Africa.

v) Establish a national coordination structure for evaluation activities in each country.

vi) Develop networking and exchanges between practitioners.

vii) Organize conferences, symposia, and seminars on current issues in evaluation with the objectives of informing, raising awareness, and harmonizing practices.

viii) Disseminate results.

### Dr. El Hadji Gueye's Expertise

Dr. El Hadji Gueye's expertise includes methods of impact evaluation for programs and projects, M&E for health expenditures, management based results, and economic evaluation for programs. He has provided consulting services for the World Bank, UNICEF, UNAIDS, UNDP and technical assistance on data collection and analysis for evaluation and management based on the results with Marginal Budgeting for Bottlenecks tool in Senegal, Burkina Faso, and Gabon.

He has coordinated and facilitated workshops on National Health Account, M&E for health programs, and strategic planning for HIV programs. He is the head of the Economics Department and the coordinator of CLEAR Initiative for Francophone Africa at the Centre Africa in d’Etudes Superieures en Gestion (CESAG) in Dakar, Senegal.

He has a PhD in econometrics and statistics, a master in econometrics, and a bachelor in quantitative methods at the faculty of Economics and Social Sciences, University of Geneva, and a bachelor in Mathematics from UCAD in Senegal.
MONITORING AND EVALUATION (M&E) have long been promoted by the development community as necessary for the enhancement of results-oriented public management, to foster accountability, and to enable evidence-based policy making to promote development effectiveness. In the developed world, this trend resulted in the adoption of new public management concepts and principles in the process of administrative reform and modernization. In the developing world, however, this agenda has been more forcefully promoted by the demand for accountability of donors and external institutions.

Apart from the obvious differences between the policy systems and socioeconomic contexts of countries with varying levels of development and democratic consolidation, it is worth noting that the external push for M&E has left an important imprint in the scope and trajectory of the adoption of these instruments in the global south. The low level of adoption and citizen engagement in the push for M&E and the absence of strong, home-grown communities of practice and attentive publics committed to this agenda have severely limited the scope for the incorporation of the logic of evaluation—and results-oriented management in general—in domestic policy circles beyond the lifecycle of specific development projects. This is, in part, because in the past decade, we have witnessed an important shift from the traditional emphasis on program evaluation, anchored in specific projects and sectors, to the central role of the creation of national evaluation capacities and the development of national evaluation systems that are more clearly anchored in domestic dynamics; local adoption; and, most important, the presence of the real economic and political conditions needed for sustainability and use for of evaluation findings in decision making, across policy areas and beyond specific government tenures.

1 CLEAR (Regional Centers for Learning on Evaluation and Results) www.clear-la.cide.edu
But what exactly are national evaluation capacities? How are they related to the construction of national evaluation systems? And, most important, why should we care about this incredibly complex endeavor?

In this article I offer an overview of the role of national evaluation systems in democratic governance and reflect on the windows of opportunity and the common challenges faced by Latin America, where relatively new democratic regimes in very unequal settings struggle to meet demands for accountability and policy efficacy. I draw on an epistemic concept of public policies and present the functional and normative link of M&E systems with democratic legitimacy and government efficacy, and then present some examples of how this link can create a virtuous circle of policy reform, as exemplified by recent developments in Latin America. Finally, I reflect upon the core challenges that must be met for these systems to be consolidated and truly become sustainable features of public management and democratic governance.

What’s Democracy Got to Do With it?

The development of national evaluation systems and the adoption of technical tools to assess program and government performance can be mistakenly associated with a technocratic push for placing political decisions in hands of a narrow group of experts, or the backdoor for overt policy influence by external actors. In my view, national evaluation systems represent precisely the opposite: they place better tools and knowledge in the hands of democratic decision making, providing an informational public good for accountability and citizen empowerment and, ultimately, securing the legitimacy of democratic regimes.

In a democratic system, citizens delegate sovereign powers to elected officials, who receive a mandate to promote substantive outcomes for the community and a series of instruments—public funds and powers, authority, regulation, administrative structures, and so forth—to advance them. Originally, the preservation of this mandate could theoretically be secured by the electoral incentive (Downs 1957) and the rule of law of the Weberian bureaucracy (O’Donnell 2010). Governments were directly accountable to the democratic will because early democratic regimes where anchored in liberal states with narrow tasks, primarily preservation of territory and the protection of private property (Gorski 2003).

During the twentieth century, in contrast, governments increasingly took direct responsibility over social realms that required a more complex production function: direct investment for productive activities, strengthening of domestic markets through regulations and subsidies, the expansion of social services, and a series of policies and programs aimed at responding to pressing social demands in several spheres. Contemporary governments are responsible for a number of public goods and services that are essential for the livelihood and welfare of the population, either directly or through their influence on the structure of opportunities for economic development and the creation of public value.

Every day, decisions on the design, distribution, and models of implementation of public programs and services affect our lives. These include
the quality and relevance of public services we receive; the structure of access to social rights such as health and education; the degree of (in-)security that prevails in our public spaces; and, in general, the structure of incentives that households, markets, and governments face. These services have an important influence on our collective ability to meet basic needs, promote development, and cope with social problems of various kinds. For this reason, understanding the effects of public policies, and the attributes that are associated with their success or failure, as well as the cost structure and functionality of the management model that characterizes them, is essential for decision making and democratic governance—and it should concern us all as citizens.

Furthermore, the scope of public responsibility has gone well beyond prototypical bread-and-butter concerns. It increasingly involves more elusive, yet equally instrumental (for the development of human capacities and functioning; Sen 1999), social objectives: future generations (sustainable development), gender, equity, the protection of intangibles, the protection of privacy, human dignity, and third-generation human rights, among others. These aspects, closely related to government capacities and performance in core social areas, have an important influence on our collective ability to meet basic needs, promote development, and cope with social problems.

Nevertheless, there are important obstacles that must be overcome for this ideal democratic equation of accountability to materialize in everyday policy debates. First, there is a collective action problem. Although citizens may be victims of poor services and poor government performance, it is not easy for them to organize and successfully mobilize to voice their specific demands and needs. In very unequal countries, more affluent citizens tend to “vote with their feet” (Tiebout 1956) and opt out of public services and their provision, thereby undermining the scope for the articulation of bottom-up social demands.

The policy literature repeatedly tells us that programs with poor or disempowered constituencies gradually become more fragile. Their quality is low and they are poorly financed because key stakeholders lack the time, resources, and information to push for quality and efficacy.

In other words, we need to track whether the theories of change implicit in any intervention and the consequent mobilization of public funds to tackle development challenges are progressing toward producing the expected result.

In addition, on the supply side, decision makers face multiple incentives and restrictions in adopting the logic of results-oriented management. They can be budget-maximizers that compete for budgets with the logic of bureaucratic control and power (Niskanen 1971); they can opt for short-term solutions that provide immediate political benefits instead of securing long-term investments that maximize public value; they can invest public resources in the construction of clientelistic and corporatist networks for
political gain; and they can also, just like citizens, make profound mistakes in their decision making (Kahneman and Tversky 2000).

Furthermore, what we should expect and demand as citizens are policy outcomes, not simply public programs. Decision makers and government agencies produce policy outputs, not necessarily linked to expected policy outcomes. For this reason, the link between government action and the transformation of the core variables around a public problem cannot be held as axiomatic. As lawyers would say, the burden of proof is a public responsibility of governments, because citizens cannot systematically give the benefit of doubt to politicians, without strings attached, and because the opportunity costs of ill-conceived and poorly implemented public interventions that utilize scarce public resources are too high.

Citizens should not accept the presumption of efficacy or give the benefit of the doubt when it comes to assessing the contribution of a public policy to achieving the social objectives that gave rise to it. And decision makers should not be entitled to make these choices by mere intuition or good intentions, in an evidence-free zone. The actual existence and verification of the instrumental-causal link between public choices and public outcomes is an inescapable public responsibility for democratic governance. But this link, an anchor of results-oriented management, needs to be constructed. It requires entire administrative and information management processes that speak the language of results and feed into bureaucratic routines and political language.

As Luis F. Aguilar (2010) has suggested, the great challenge of the democracies of the twenty-first century is not so much “who governs” but “how she governs,” based on what reasons and arguments; through which instruments; and, above all, with what results. This is a core mandate of democratic governance in the twenty-first century. Without credible performance and efficacy, democratic legitimacy is also at stake.

An Epistemic View: The Importance of Speaking Truth to Power

Hugh Heclo used to say that public policies are a way to puzzle on society’s behalf. So conceived, public policy is a tentative answer—and collective commitment—to solve public problems and pursue socially desirable goals. If we consider that public policies are theories or models for social transformation—hypotheses for solving complex public problems—and they are implemented by public organizations in constant interaction with an environment that moves much faster than governments anywhere in the world, the enormous complexity and uncertainty in which the simplest of interventions operates requires us to create reliable and timely tools to track the adequacy and validity of those hypotheses.

In other words, we need to track whether the theories of change implicit in any intervention and the consequent mobilization of public funds to tackle development challenges are progressing toward producing the expected result. In that sense, the role of evaluation systems in inextricably linked to that of knowledge generation, dissemination, and learning as a routine task.
of modern public administration. So conceived, M&E allows for the creation of a learning curve in which consciously exploratory trial and error in policy making can promote better programs, better choices, and better results for the citizenry.

While it is true that making mistakes is as human as it is inevitable, it is also true that making mistakes in key decisions for the public interest (with public funds) entails serious moral and political implications. In addition to the distributional consequences and opportunity costs of erratic public programs that lack a clear roadmap and an explicit commitment for the transformation of specific social indicators, the erosion of the effectiveness and responsiveness of the state is a major danger to the legitimacy and sustainability of democratic governments. This is why the M&E agenda as a new language of public administration will allow governments to better serve citizens: it will reduce information asymmetries between rulers and citizens, ensure that fiscal resources are channeled in reasoned and reasonable ways for investment, and assure that public programs meet clearly identified priorities and are developed with a high probability of positively affecting the quality of life of the population.

Do programs work? How do we know whether programs work? Are program diagnostics accurate? Are public interventions reaching the population they are expected to serve? Is the quality of services adequate? Are there alternative ways to advance certain goals, at a lower burden to public budgets? How does one decide between complex trade-offs in the allocation of scarce resources? How can we collectively learn from experience and that knowledge to enhance the effects of government interventions? Is there a way to maximize the bang for the buck?

National evaluation systems can answer these sorts of questions. M&E systems can address the issues, from the smallest detail of project management to broad strategic discussions about intervention strategies and the great debates about the development model we want and how we can achieve it. In the countries of the region, we have made major progress in the basic institutional framework for assessing, monitoring, and managing for results. However, as Aaron Wildavsky once noted, policy analysis and evaluation are a way of speaking truth to power. It is thus understandable that M&E systems can face significant internal resistance. They can even be perceived as an organizational threat to the public sector by demanding transparency and greater accountability. To the extent that undergoing public scrutiny and evaluation can undermine assurances given through bureaucratic routines, operating logics, and decision making behind closed doors, it is natural that these systems have generated resistance.

For this reason, the creation of a national evaluation system might be necessary, yet insufficient on its own to create, promote, and enhance the national evaluation capacities that are required

The link between institutional frameworks (national evaluation systems) and actual capacity for evaluation has been at the heart of institution-building theory and debates.
to stay the course and promote the utilization of evaluation results by different audiences and for several interrelated purposes. These purposed include budget control, program improvement, accountability, policy reform, transparency, and the like. For M&E systems to fulfill their mission, the public policy debate must be enriched with better arguments, evidence, and diverse sources of information that are credible and widely discussed in the public arena.

M&E systems are information systems and processes that are anchored in results and evidence-based inputs. They allow both citizens and governments to know the truth about the effectiveness and efficiency of policies, to find out whether public actions, instruments, and strategies are correct or whether basic premises and assumptions should be revised. That is the primary role of M&E systems.

National Evaluation Systems and the Quest for Evaluation Capacities

The link between institutional frameworks (national evaluation systems) and actual capacity for evaluation has been at the heart of institution-building theory and debates. Should existing baseline capacities precede the enactment of formal systems for evaluation? Are they a direct result of institutional frameworks, or do they always interact in far from linear relationships?

Instead of trying to solve that puzzle, I will propose some lessons from recent Latin American trajectories in the development of national evaluation systems. In Latin America, the third wave of democratization created conditions for a rising demand for accountability, transparency, and scope for citizen participation in public decisions. Countries such as Chile, Colombia, and Mexico stand out for having developed strong institutional frameworks aimed at institutionalizing M&E to promote results-oriented budgeting and better government performance. Although other countries in the region have followed suit with differing pace and scope, Latin America can be described as a region in which the background conditions and the political context have been conducive to the advancement of the M&E agenda. Within the next decade of so, it is reasonable to expect that all countries, with the likely exceptions of Venezuela and Cuba, will have some sort of national evaluation systems in place, as incipient as it might be.

However, this important step will not be sufficient to secure strong and reliable evaluation capacities across policy sectors and bureaucratic levels, and extending beyond governmental actors and small groups of stakeholders. As mentioned above, M&E systems are a first step for the production for public information that is reliable, relevant, and results-oriented. But effective capacities for evaluation require system-level changes and positive feedback and learning cycles, and they depend on factors that are external to the evaluation systems.

First, evaluations have to be independent and credible, and they must be reliable as impartial sources of information. For this to be the case, there has to be a critical mass or community of practice willing and able to provide external evaluation services, and, on the demand side,
agents commissioning and overseeing evaluation must be well-versed in public administration and be able to identify key evaluation questions, core methodological requirements, and draft terms of reference and planned budgets accordingly.

For these conditions to develop, there must be a clear external demand for results-oriented information that is channeled into the political and policy process with some degree of saliency. If both commissioners and evaluators expect the product of their evaluation efforts to be forgotten, neglected, or ignored, they will have no reason to invest in the quality and credibility of evaluation. Furthermore, if the decision makers responsible for key performance and process indicators of the programs to be evaluated do not expect evaluation results to resonate within the bureaucratic and/or political arenas, there is no reason for them target their efforts at inducing better outcomes that can be identified and recognized through M&E. In turn, citizens and attentive publics who are not informed and aware of the potential role of evaluation results and monitoring data will not understand their link to substantive development promotion. Without such understanding, they will fail to utilize these results in their push for accountability and better government responses to public problems.

With this description, I am not trying to suggest that evaluation capacity can only be developed when public servants, government, and citizens are thoroughly informed, highly professional, and spirited and strong advocates of evaluation. What I am suggesting is that actual capacity development can be understood as an endogenous result of the internal and external interactions enabled by national evaluation systems, but that strongly depend on other background conditions: the push-and-pull factors for the use of evaluation in the standard politics of political pluralism (democratic governance)

At the CLEAR Center for Latin America, we have a mandate to promote capacity building for evaluation, monitoring, and performance-based management to promote development at the regional level, as well as to promote global cross-dissemination of good practices and learning spaces related to our core functions (technical assistance, knowledge generation, and training). In order to fulfill our mission, we have adopted a systemic view of the role of capacity-building like the one proposed above: we seek to promote the understanding, adoption, and mobilization of M&E information as a new language of the democratic political interaction that is already taking place in pluralist systems, yet does not usually occur in the language of results and evidence-based policy making. To lower barriers of entry to this language, we promote training, awareness, and development of technical expertise among wider audiences and the diverse stakeholders who already occupy center stage in policy debates—civil society, activists, watchdog organizations, and government agencies at the national and subnational levels—and seek to foster leadership among key stakeholders in the promotion of this agenda.

By providing accessible and reliable information from the region, we seek to promote knowledge dissemination and accelerate learning processes among peer nations. Furthermore, by mobilizing existing information from national evaluation systems and spreading it beyond traditional audiences, we aim to promote use and adoption
in order to inform policy debates and highlight the importance of M&E for good governance.

In Latin America, the diffusion of conditional cash transfers for poverty alleviation clearly has been influenced by the role of credible impact evaluation and has already resulted in successful poverty alleviation interventions and the protection of otherwise neglected constituencies, such as the rural poor and indigenous populations. In the health sector, information and cross-learning has promoted the development of aggressive public programs to prevent catastrophic expenses from out-of-pocket spending on health care. In the field of financial deepening and microcredit, policy debates have also been anchored in rigorous evaluation and dynamic learning processes. Food security and education are other areas in which the information from evaluation systems has been pivotal in promoting policy reform and accountability. Although national evaluation systems and capacity for evaluation are far from consolidated, I believe that there is window of opportunity to promote this agenda and further advance our common goal of development.

**References**


---

Dr. Claudia Maldonado is currently Professor-Researcher at the Public Administration Division of the Centro de Investigación y Docencia Económicas (Center for Research and Teaching in Economics) a Mexican center for research and higher education specialized in the social sciences. Her research focuses on the political economy of conditional cash transfers in Latin America and comparative development (Mexico and Brazil), social policy and program evaluation.

Prior to her current position, Dr. Maldonado coordinated M&E capacity-building efforts for Mexico’s Consejo Nacional de Evaluación de la Política de Desarrollo Social (National Council for the Evaluation of Social Development Policy, CONEVAL) and worked in the General Direction for Evaluation at Oportunidades, the largest conditional cash transfer program in Mexico.

Dr. Maldonado earned her Ph.D in Political Science from the University of Notre Dame, and holds a Master’s Degree in Public Affairs from Princeton University.
Mohamed Manai, Manager, Operations Evaluation Department, AfDB

CONSIDERABLE PROGRESS HAS been made in M&E over the last 20 years. During this period, the United Nations Development Programme (UNDP) has played a key role in strengthening M&E capacity and bolstering the role of other funders, including that of the African Development Bank.

A 1990 regional conference, held under the auspices of the Organisation for Economic Co-operation and Development (OECD) (followed by other regional conferences in Kuala Lumpur, Malaysia, and in Quito, Ecuador) highlighted the notable progress made in M&E. In particular, discussions underscored that the trend observed—continued to establish good governance and effective management of public resources—to influence governments and donor relationships to build an M&E culture. The conferences also highlighted that countries are not all at the same level of development of national capacity assessment and that each country has followed its own path. This calls for an approach tailored to the political, administrative, and cultural context and the level of socioeconomic development of each country. It also calls for a preliminary diagnosis of both supply and demand for M&E to identify ways and means of implementing the institutionalization, professionalization, and implementation of M&E, whether in a centralized, decentralized, or participatory framework or other model.

Any national strategy for evaluation capacity building must emanate from the country itself through consultation with all stakeholders regarding national development plans or strategies for poverty alleviation and growth. Donors must coordinate their actions based on the resulting guidelines and national choices and work together with governments.
There are therefore no best practices, but there are good practices. All routes are possible if they are oriented toward improving the legislative framework for evaluation; the administrative framework, "constitutionalization," or regulatory framework; or ensuring the involvement and active participation of civil society.

Any national strategy for evaluation capacity building must emanate from the country itself through consultation with all stakeholders regarding national development plans or strategies for poverty alleviation and growth. Donors must coordinate their actions based on the resulting guidelines and national choices and work together with governments.

A good example of this is the case of the Democratic Republic of Congo (DRC), where the government sought support from the African Development Bank (AfDB), the World Bank, and the UNDP to modernize public administration. These institutions were engaged to act together in a coordinated framework. The UNDP completed the diagnostic study of national capacity, while the AfDB and the World Bank set up a framework for joint implementation of results-based management system assistance. Such organization of tasks avoids duplication and overlap of efforts. The CLEAR (the Regional Centers for Learning on Evaluation and Results) initiative, promoted by the World Bank and supported by the AfDB and other donors, will build on the diagnosis of national capacity assessments carried out by the UNDP. It will not reinvent the wheel, but, instead, build on existing achievements.

The following two proposals can help move forward the NECD agenda:

1. **Create a portal on the development of national capacity assessment.** This portal could be hosted on the UNDP website (or another suitable site), and would provide access to the experiences of countries that wish to share their best practices and their diagnostic methodological tools for national capacity assessment. This portal would also allow access to decision makers, planners, researchers, academics, practitioners, and national and regional networks. It would constitute a framework for exchange and inspiration through forum discussions on assessments of capacity building. It would provide links to sites such as those of the OECD, World Bank, and UNDP.

2. **Encourage donors to work toward the convergence, harmonization, and strengthening of policies in order to avoid duplication or overlap.** The harmonization framework is available through the application of the principles of the Paris Declaration and the Accra Agenda for Action. It is timely to recall that it is possible to develop cooperative support strategies. The CLEAR initiative, for example, is a multidonor program that has created regional centers in Asia, Latin America, and Africa (Francophone and Anglophone), but it may evolve into an expanded program to meet the large volume of requests for evaluation capacity building. This program will fund not only diagnostic studies but also trade activities, travel studies, training, and knowledge management. We hope that this program can interest other donors and evolve quickly.

In sum, coordination of donors in an effective partnership remains in the hands of the recipient countries, with shared responsibility and mutual accountability for improved policies and efficiency of development aid.
Rebirth, Restoration, Reclamation, and Responsibilities of the Evaluation Function of Africa

Dr. Sukai Prom-Jackson, Inspector, Joint Inspection Unit, United Nations

Conceptual Framework for Evaluation and National Evaluation Capacity Development¹

Why Africa Has Not Developed to Its Full Potential
Moussa Seye (African savant) on challenges to African growth and development
(August 7, 2013) (translated from Wolof)

—If someone lends you his eyes, you will only see what he wants you to see.
The eyes are the windows to the mind. They will direct what you think.

—If someone lends you his language, you will use the models he has developed to express reality based on his own ecological and historical background.
There is no sovereign that has developed on the language of another culture.

—If someone lends you his mind based on how he defines value and truth, you will work for it and continue to fix it to fit.
You will never know and you will never become a fit human being.

—If you do not farm what you eat, and you do not eat what you farm, you will never gain control of your appetite and stomach.

¹ This note is a follow-up to reflections afforded by the opportunity to participate in the 2012 African Thoughts Leaders Forum in Bellagio, organized by the Center for Learning on Evaluation and Results. It incorporates dialogues, reflections, and notes based on consultations with several thought leaders in Africa. It provides the conceptual framework that should guide the development of the evaluation function in Africa. Follow-up notes include using the framework to analyze: (i) the state of affairs in Africa, (ii) support for evaluation capacity development by the United Nations systems and development partners, and (iii) what is needed by nationals to move the process and engage all pertinent national institutions involved in critical thinking and action.
Introduction

Moussa Seye’s words embody lessons about the underlying problems of development in Africa. They also apply to evaluation and national evaluation capacity development.

Evaluation and support for evaluation capacity development should avoid being part of the problem of the development of Africa by using lenses, language, models, and imported best-practice models that are not home-grown. To play a constructive role, the theory and practice of evaluation in Africa and support for national evaluation capacity development must:

- Understand the added value of evaluation for Africa.
- Be consciously sensitive and take inspiration from inherited legacies.
- Confront and direct the current varying, often hybrid, and complex realities of the present.
- Build the futures and securities for Africa in a global context.

Evaluation has to be grounded in a clear understanding of the goals being articulated for the rebirth, restoration, reclamation, and responsibilities of development for Africa. These goals present an alternative one illustrated Moussa Seye’s words.

This paper presents the core capacities required for evaluation and a conceptual framework to support national capacity development that is founded on principles for the rebirth, restoration, reclamation, and responsibilities of evaluation for Africa. The paper covers: evaluation and its fundamental capacities, capacities for evaluation required for an effective role in the development of Africa, and a conceptual framework for national evaluation capacity development. Follow-up work is planned for using this framework to assess what is happening with evaluation in Africa and the capacity development support provided by development partners.

Evaluation and Capacities for Critical Thought

Evaluation is judgment of value or worth and provides information to support decision making. In development evaluation, it supports accountability for the effective use of resources, lessons for improvement, knowledge sharing, and the distillation of this knowledge for use within organizations and in the wider global community.

The main factor that makes evaluation stand out from other oversight functions in organizations is the critical mode of thinking it requires. It is a rigorous process of inquiry. It examines actions and results and asks the questions: Are we doing the right things? Are we doing things right? Are we getting results that make a difference? Are these the right results, and what is the impact and value? What explains the difference made and what enhances impact and sustainability?

In providing answers to these questions, evaluation uses the systematic methods and
tools of basic scientific inquiry. As such, it draws on critical reasoning skills and applies both divergent and convergent models of reasoning and analysis. It also uses principles from management science to enhance utility. It thus abides by a set of defined norms and standards—based on a valuation system—that seek to enhance the development of strong evidence that is credible and trustworthy (that is, impartial, valid, and reliable), and the use of such evidence and recommendations from the evaluation.

To enhance credibility and utility in supporting decision-making processes associated with accountability, improvement, learning, and knowledge management, governments and organizations have put in place evaluation functions with defined structures, institutional frameworks, and systems and mechanisms. Figure 1 provides an illustration of the evaluation function using the key elements of demand, supply, and results and their respective capabilities as a framework.²

**Evaluation Capacity Development and Role of Evaluation for the Development of Africa**

*National evaluation capacity development* is an *endogenous national process* to develop the capacity for “evaluative thinking.” This encompasses the use critical thinking and divergent and convergent reasoning to produce credible (impartial, valid, and reliable) and trustworthy evidence to support (i) decision making and (ii) action that would benefit national growth and development, including the development of social and intellectual capital. This national process is carried out at all levels of society and is institutionalized, in particular as part of the business model for education and schooling (formal and informal), government and national governance, universities, think tanks and research centers, the academic community, and the portions of the private sector involved in innovation and development.

This capacity for evaluative thinking is made more urgent as the people of Africa confront the current and emergent global challenges, for which existing blueprints are not always adequate, and the development challenges of Africa tied to Africa’s need to take a leadership role in decision making and action with conscious reflection, resilience, and adaptability to address six key framing conditions, outlined in Box 1.

² The framework illustrated was developed by the Joint Inspection Unit of the UN system. See “Analysis of the Evaluation Function of the UN System” (2013). It builds on previous work by the author as staff of the World Bank.
Figure 1: The Evaluation Function—Key Components and Systems Framework for National Evaluation Capacity Development

Demand and Intentionality for Evaluation—REBIRTH
- National context and drivers of demand
- Nature and level of demand
- Vision and strategy for evaluation

Supply, Adaptation, Growth of Function—RESTORATION

Results of Function Utility and Impact RECLAMATION

ENABLING ENVIRONMENT: NATIONAL, ORGANIZATION AL & INSTITUTIONAL EVALUATION FRAMEWORK
- Governance; legislative, executive, management
- Evaluation architecture and structures at all levels
- Institutional framework: Policy, Norms,

RELEVANCE, RESPONSIVENESS, EFFICIENCY AND ADAPTABILITY
- Coverage
- Responsiveness challenges
- Efficiency
- Adaptability of function for continuous improvement and growth

CREDIBILITY: IMPARTIALITY AND BALANCED PERSPECTIVES
- Structural Independence
- Functional Independence
- Controls against management influence
- Professional/Technical Independence
- Behavioral Independence

CONDITIONS IN PLACE TO ENHANCE USE
- Report Quality
- Dissemination and communication strategy
- Timeliness
- Accessibility
- Transparency
- Follow-up mechanisms

NATURE AND LEVEL OF USE
- Recommendation implementation rates
- Use of strategic decision-making
- Corporate use
- Formative Use and Improvement

EFFECT OF USE
- Effect of use on changes or development
- Value for money
- Mutual accountability
- Learning nations and communities
- Contribution to knowledge for development and evaluation

Box 1: Six major framing conditions for evaluation in Africa

- Africa’s long-term growth prospects are strong, propelled by both external trends in the global economy and internal changes in the continent’s societies and economies.
- Resource wealth and the responsibility for managing this wealth in the interest of Africa.
- The shift in the global economic center of gravity, the role of emergent global partners, and the responsibility to remain strong and draw from lessons to engage new dominant partners, such as China.
- The wave of democratization and citizen activism within a context of a dominant and growing youth population, high rates of employment, and extensive open communication systems.
- The eventualities of the future as affected by globalization, climate change, economic and political instability, and demands for rapid response at all levels (individual, organizational, and governmental).
- The salience of technologies and innovation and the role of home-grown technologies and methods in the formulation of appropriate models that are “Africa-centered, driven by African interests, and rooted in African frames of reference.”

Sources: Zenda and Rob Moore; Roxburgh and others 2010; Jeng 2012.
The first priority of this approach must be **rebirth**. The meaning of rebirth may vary, but what it entails is a break from the past and its disempowering features. This is necessary to allow space for ideas, practices, and routines that are not laden with negative legacies.

To enhance its role in Africa as a critical instrument to address these emerging changes and imperatives, evaluation must seek to not be part of the problem of development in Africa, but rather to provide solutions to problems and contribute to the future of growth and development and the ultimate goals of dignity, peace, and prosperity.

Approaches to addressing these framing conditions may vary. It is submitted that for any suggested framework—including those for evaluation—to serve as a critical tool or to outline the capacities needed to guide national development, it must have the following four elements for a home-grown framework of action. These are referred to in the rest of the paper as the **four Rs** (4Rs)

**Rebirth:** Africa’s experiences since independence suggest that development and evaluation techniques, theories, and practices have been a central part of Africa’s problems. They cannot therefore continue to be the sole frameworks upon which solutions are offered. A new home-grown approach specific to Africa is needed.

The first priority of this approach must be **rebirth**. The meaning of rebirth may vary, but what it entails is a break from the past and its disempowering features. This is necessary to allow space for ideas, practices, and routines that are not laden with negative legacies. However, for there to be a meaningful break, rebirth must be total, purposeful, and home-grown. It is given expression in the French phrase: enracinement puis ouverture.

**Restoration:** Africa’s encounters have exposed and subjected it to agonies, violation, and depersonalization. There is an urgent need for restoration: restoration of pride; restoration of heritage in development narratives; restoration of ambition; and restoration of genuine leadership. But it is important to note that restoration is not about norms-localization or introducing foreign models. Rather, it is about putting back the missing links that were compromised by the dominance of Eurocentrism in Africa.

**Reclamation:** The fault lines in current development models and evaluation methods mean that Africa must now turn to and explore home-grown knowledge systems. This is necessary for the reclamation of African spaces and ownership in development narratives. A potential benefit is the possibility of seeing Africa through a totally new lens that adds pride, confidence, and relevance to home-grown knowledge systems. Achieving this requires conceiving methods and models that are particular to, and reflective of, African sensibilities.
Responsibility: In rebirth, restoration, and reclamation, there is potential in the formulation of an Africa-centered, Africa-specific, and Africa-driven evaluation function and models for evaluation. These three Rs also provide the framework for national evaluation capacity development. How to make this happen is a critical consideration. It requires leadership and action from Africans. What is different this time in national ownership is the nature of the vision that would drive the process—one that is focused on the interests of Africa. This vision draws on the legacies and considers rebirth, restoration, and reclamation with due regard to the vastness of Africa, the interlinkages with various cultures, the current contexts, and the sustainability of the progress that has been made.

Drawing from this background on the 4Rs, the following passage provides guiding principles that must be taken into consideration in the conceptualization and development of the evaluation function and in strengthening national capacities for evaluation in Africa.

Conceptual Framework for the Evaluation Function and Capacities for Evaluation in Africa

Demand, supply, results for accountability, improvement, and knowledge management within a well-defined, integrated framework that is “Africa-centered”

Definition: Evaluation is a judgment of the value or worth of an entity. The role of evaluation for development of Africa must include defining what is valued for moving Africa forward. Evaluation is a systematic process based on scientific and management principles. The process of generating evidence and knowledge must respect ways of knowing in Africa, including the importance of active involvement in the construction, management, and use of knowledge.

Systems model—the national evaluation system: Given the role of evaluation for accountability, improvement, learning, and knowledge management in the interest of the nation, the national evaluation function and capacities required must necessarily use an integrated and holistic framework that applies a systems model. This includes:

1. Acknowledging all key national players and institutions involved in the definition of, demand for, supply of, and use of evaluation. This includes nationals and civil societies, government, education and academic institutions (involved in production of human resources, knowledge, and intellectual capital), research and analytic centers, and think tanks involved in knowledge production (see Figure 2).

2. It also includes understanding the interrelationships among demand, supply and use between the national system and the regional and international systems. Figure 2 provides an illustration of these interrelationships.
Guiding principles for the evaluation function: Rebirth, Restoration, Reclamation, and Responsibility: The following set of guiding principles from the 4Rs provide the framework that must be part of the constituent elements in defining the key elements of the function of national systems and that should guide capacity development and support. The 4Rs are interrelated. However, the way they fit with the key elements of the evaluation function for demand, supply, and results is illustrated in Figure 1.

Rebirth: Defining demand, purpose, and added value of evaluation. The vision for evaluation in Africa must promote evaluation that informs development that is Africa-centered and rooted in African realities and the issues that are critical for moving Africa into the future.

- In so doing, evaluation must move beyond current predominant models of knowledge generation built on hypothesis testing of imported “best practice.” It has to value various forms of knowing recognized in African societies as well as in various parts of the world. Its business model must be inclusive in setting an agenda of critical issues important for the future of Africa, addressing key challenges and all aspects of society and all groups, beyond those pertinent to development partners.

- Given that there are no blueprints, evaluation must focus on developing understandings of both the processes that inform improvement in a meaningful way and the end results. In this regard, formative evaluation takes center stage along with summative and accountability evaluation.

- The purpose of doing evaluation for accountability must depart from a focus on accountability to external donors and move to the accountability of the government to its citizens, as well as accountability of citizens to each other.
A quarterly knowledge publication of the Operations Evaluation Department of the African Development Bank Group

- A paramount goal of evaluation is to build dynamic learning groups and learning nations that are resilient and adaptable to fast-moving change and focused on the use of knowledge for innovation and change.

- Efforts directed at knowledge management must focus on garnering knowledge in Africa to direct and manage external support and to contribute to global knowledge and platforms.

- Evaluation must acknowledge that in a world of rapid change, knowledge is a flow. It moves and changes. The value of any one knowledge product depends on how it is developed, used, and subjected to innovation. It has value when used.

  **Restoration:** *Supply guided by restoration of knowledge based on the African lens and philosophy and used to define norms, standards, models, methods, and sensitivity to context:*

- Move from marginalization to the mainstream process of knowledge construction. Evaluation must take African voices into the mainstream of development.

- The function needs to partake of African indigenous philosophies and ways of knowing and learning. This has implications for approaches and methods. Evaluation for Africa must be built to incorporate indigenous ways of knowledge generation, judgment, and use.

- Evaluation must be poised to develop norms and standards built on African value systems in their pure or syncretic form, as deemed best for the future of Africa.

- Given diversity and varying contexts and cultures, it must incorporate traditional values as well as values in the nexus between Africa and Europe, Africa and Islam, Africa and complexity, and Africa in a modern world of fast-moving information technology.

- It must have space for the independence to make judgments. To do this, one must return to tradition—refinement must take place first in one’s culture, and then in the world.

  **Reclamation and Responsibility for Results:**
  Evaluation must assume leadership for Africa and take its place in defining high-impact evaluation functions that respond to varying demands in a balanced way. It must reclaim its position in the construction of knowledge and in serving as a legitimate agent of change. It should seek to develop a vision for evaluation, with key elements that include balancing the fourfold purpose of evaluation for accountability to nations and to development partners, improvement of practice, building learning nations and enhancing innovations in society, and contributing to global knowledge and to the field of evaluation.

- The vision should not limit evaluation to activities by government. It should highlight the role of national institutions involved in the supply, demand, and use of evaluation, including government, education and schooling, think tanks and research and analytic centers, universities and the academic community, the private sector, and nationals and civil societies at large.

- It should seek to enhance the development of a learning nation engaged in constant reflection, critical reasoning, action, and
innovation. This requires building resilience and adaptability in responding to rapid change on an informed basis and on a solid ground for development effectiveness and sustainability.

- It has a responsibility to enhance cognitive diversity in the active construction of knowledge. This requires an interdisciplinary approach that brings in various competences and engages the global debate with valid evidence.

- It will acknowledge complexities of development and the hybrid nature of various contexts and development progress made in Africa. This will complement emerging methods and models in evaluation, including complexity science, adaptive model, and models of cognitive diversity.

- The role of the evaluator is not that of a reporter. Rather the evaluator is a central change agent who creates room for dialogue, engages in the rigorous exchange of knowledge, and in creating possibilities and solutions.

- It must recognize and take a lead role in applying more integrated and systems models for evaluating the development program of partners. This includes taking a lead role to define what work is conducted and to coordinate the evaluation agendas of development partners in alignment with the national evaluation agenda.

- It should help enhance understanding of the complexity of development and how traditional evaluations of development by partners are detrimental to development impact. Besides the nature of the models, the business model of "silo evaluations" by different partners is equally destructive and kills development: "When you cut an elephant into 17 pieces to manage it better, you don’t get 17 smaller more manageable elephants. You get a dead, chopped-up elephants." The large errors involved in separate evaluations on the same phenomena raise questions of the validity of evaluations by development partners.

- Evaluation should consider an open-systems model, given the interdependence and nature of the change process.  

---


5 Key features of the open systems for evaluation capacity development include:

(i) Understanding the subsystems operating at various levels of governance structure.

(ii) Emphasizing both the internal and the external the environments in which the function operates, since each plays an equally important role for quality, impact, and sustainability.

(iii) Sometimes subsystems relate to each other in a logical manner (inputs, activities, outputs, outcomes, and impact), but at other times changes in one level are likely to have nonlinear/causal and often unpredictable or unintended effects.

(iv) This is because change is a process and it is people who change. The change process goes through various stages of maturity or development in adopting guiding principles and set norms and standards due to a number of factors that are embedded in the function itself—organizational climate, nature of demand and its evolution, characteristics of decision makers, leadership and champions, organizational culture and its absorptive capacity for risk, learning, improvement or innovation, external global conditions—all have a role in driving the change process toward a high-quality and high-impact evaluation function.

(v) Performance or level of development in these components and attributes follows a process of growth and maturation, from reorienting, to mechanical use, to a routine level, to a level of concern for impact and the sustainability of the function.

(vi) There are ideal norms and standards setting optimal rules of the game. There is generally no perfect configuration of subsystems that should be used as an ideal. There is no such thing as "one size fits all." Systems are best defined as "fit for purpose" at a particular time, given history and context.
• It should take a lead role in directing substantial focus on formative evaluation to enhance understanding of what works, why, and how. This is important for testing development hypotheses and for timely engineering and prevention of the notion of development as a scientific laboratory. The complex operations in complicated contexts and infrastructure for evaluation must focus on continuous monitoring, assessment, and redirection.

References


Sukai Prom-Jackson is an Inspector with the Joint Inspection Unit of the UN System. She has over 20 years of professional experience and leadership in the management and conduct of research and evaluation, in policy formulation and strategic planning and program development, and in the facilitation of learning.

She has worked for the UNDP, World Bank, U.S. government, universities, and research centers in the United States. She has a PhD in educational psychology (1982) with focus on policy research, measurement, and evaluation from Howard University (U.S.).
DEVELOPING AN EVALUATION culture within the Bank and in regional member countries is one of the cornerstones of the independent evaluation policy recently submitted for approval to the Committee on Operations and Development Effectiveness (CODE). This objective is based on the notion that promoting an evaluation culture within an institution such as the African Development Bank places an emphasis on evaluation outcomes and on the lessons learned from successes as well as failures. An institution with an evaluation culture assumes that its members:

- Agree to use evaluation outcomes and lessons learned.
- Have a sufficient understanding of why the institution uses evaluation outcomes.
- Can develop and implement auto-evaluation systems.
- Use evaluation outcomes with the intention of supporting institutional change and development.

Promoting an evaluation culture

The Bank and its stakeholders seek to promote a culture of evaluating outcomes, which can only occur if evaluation is at the center of this effort. Despite the increasing engagement in favor of outcomes already recorded within the Bank, promoting an evaluation culture presupposes a large program. The Operations Evaluation (OPEV) Department will have to collaborate with other partners in this effort and ensure that its work completes other initiatives, such as the quality assurance work done by the Bank’s Quality Assurance & Results Department (ORQR), a natural strategic partner in achieving this objective.
Reinforcing stakeholder engagement

It is important to mention that engagement occurs in three main areas:

- **Within the framework of the Bank’s project cycles.** By examining the efficiency and effectiveness of the Quality Assurance Reviews for projects, policies, and strategies during the initial phases of conception and development, and by questioning the merit of the quality standards used, OPEV offers institutional lessons and practices and brings in past experiences to reinforce these preliminary quality assurance mechanisms. Furthermore, in evaluating the policies and sector-based strategies of grouped evaluations, OPEV examines the extent to which past lessons have been taken into account in said policies or sector-based strategies in each country. Similarly, in validating Project Completion Reports, OPEV present the sector-based or regional governing bodies with the merits of the findings from the validation and the Project Completion Reports. This ensures optimal learning for staff from the relevant regional and sectoral departments.

- **Within the framework of individual evaluations.** Engaging management and relevant personnel at critical moments of the evaluation cycle (evaluation framework document, identification of stakeholder and reference groups, dissemination of knowledge from findings in the field and from interactions with beneficiaries) is a best practice that has been adopted by OPEV as part of the development and implementation of evaluations of projects and of sector- or country-based strategies. This engagement generally involves dialogue around evaluation objectives, evaluation questions, approaches, and methodologies used to judge factors of success and failure, as well as the communications and knowledge-sharing strategies.

- **Within the framework of planning OPEV’s work program.** Consultations with management and with the Committee on Operations and Development Effectiveness (CODE), which take place during the development of independent evaluation the rolling work program, seeks stakeholder buy-in for evaluation themes to increase their usefulness. Not only does this approach favor the emergence of tangible demands from decision makers and development-program planners searching for lessons and experiences, but it also fosters the development of an evaluative culture based on the use of outcomes and feedback.

OPEV’s engagement must go beyond the confines of the Bank. First, it must foster interactions between beneficiary countries by including them as much as possible in evaluations of projects (thematic, sector- or country-based); second, it needs to develop cooperation with the international evaluation community and other organizations interested in supporting evaluation in Africa.

Increasing the impact of evaluations in regional member countries

There is increasing desire for evaluation knowledge in AfDB regional member countries. Making sure that the knowledge generated by OPEV is made available to them is an indispensable step in ensuring that evaluations are impactful and useful. Within this framework, OPEV promotes
the development of an evaluation culture in three ways:

- **Regional member country participation in the evaluation process.** Participating in country strategy evaluations is encouraged but not easily achievable due to countries’ low capacities. National stakeholder involvement throughout the evaluation process must be accompanied by awareness-raising sessions on the merits of national evaluation systems, within a framework of implementing a management approach that is based on development results and the rational use of public resources.

- **Raising awareness, by targeting regional member countries in communication efforts.** Knowledge-sharing sessions, conducted in regional resource centers, facilitate participation by regional member countries. The recent workshop organized at the AfDB Southern Africa Resource Center (SARC) attests to the high level of interest of regional country representatives, who, in the case of the SARC, made substantive contributions on subjects such as private sector development, Bank assistance to regional integration, economic and sector-based reviews, and the development of general infrastructure (transportation in particular).

- **Supporting regional member country efforts to develop national evaluation capacities.** Regional member countries are also looking to generate evaluative knowledge from their own programs and development projects, whether they are financed externally or not. They are increasingly interested in reinforcing national evaluation capacities. OPEV has already begun to undertake reinforcement activities in countries where a tangible need has been demonstrated (such as Ethiopia and Tanzania), focusing its efforts on ministries in charge of planning by ensuring the practical and hands-on training of personnel in charge of post-evaluation monitoring for national development programs. In other words, OPEV will support the development of an evaluation culture in regional member countries by supporting their evaluation systems. Such systems help ensure that evaluations make a difference in policies and in practice.

OPEV is in the process of deepening its partnerships with other organizations in the international evaluation community, and it continues...
to participate in major international evaluation communities such as the evaluation network of the Development Aid Committee of the Organisation for Economic Co-operation and Development, and the Evaluation Cooperation Group. OPEV continues to develop strong and mutually beneficial relations with bilateral aid entities and engages with organizations focused on evaluation, such as the African Evaluation Association, various regional centers for learning from evaluations and outcomes, EvalPartners, the International Initiative for Evaluation Impact, and philanthropic organizations interested in supporting evaluation efforts.

Managing and disseminating knowledge

OPEV uses various mediums and technologies to foster knowledge sharing:

- **Evaluation reports**, conceived to convey key messages clearly and simply are paired with other evaluation publications (“eVALUation Matters”) and short, ad hoc memos that flag important issues for the benefit of the Board of Directors and senior management.

- **The OPEV website**, the department’s primary interface with the world has been redesigned to be more user-friendly and to allow users to quickly locate the knowledge products they need.

- **A database of lessons learned**, which will allow users to search for information by theme or by country. OPEV will collect lessons from other organizations to maximize the usefulness of the lessons database. This system will include independent evaluations as well as Completion Report Reviews, and it will cover a wide range of evaluation products and lessons.

- **Face-to-face interactions** with the Bank’s operations staff will also be reinforced. The Evaluation Community of Practice¹ will be expanded and will also target operations personnel when appropriate. OPEV will organize other knowledge events to reach a larger audience (such as management and national offices, as well as other staff involved in project and program conception, implementation, monitoring, policy, and quality assurance).

- **Regional member country awareness** will be increased through regional workshops and videoconferences, which allow for dissemination and discussion of knowledge generated by evaluation work. To further these efforts, OPEV will work in close collaboration with country teams and with the Bank’s regional resource centers and offices.

Implementing evaluation recommendations

Discussing the implications of evaluation outcomes for future policies and operations is essential for promoting an evaluation culture and ensuring that evaluation recommendations are effectively implemented. This underscores the importance of reporting and sharing knowledge and establishes a link between evaluative knowledge and sustained evaluation of development.

A mechanism to strengthen management action plans for implementing recommendations from OPEV will be launched shortly.
Such a mechanism will monitor the way management receives recommendations, as well as the clarity of its response through precise actions and initiatives spaced over time. The entities involved in reporting management actions will ensure that actions taken are transparent and that CODE is informed. It is equally crucial to monitor the implementation of evaluation recommendations to ensure the credibility and usefulness of independent evaluations.

**Promoting an evaluation culture within the Bank: constraints and challenges**

The main constraints and challenges are as follows:

- Level of support from all levels of management for auto-evaluation and the seriousness with which recommendations from independent evaluation results are implemented. The quality control process for auto-evaluation products in the Bank’s active portfolio needs to be reinforced to reduce the gap that has been noted between auto-evaluations and independent evaluation in examining portfolio performance.

- Level of support for the auditing and quality assurance bodies in a decentralized environment. Common standards of quality need to be linked to lessons from experiences or lessons that have served as the basis for the design of new policies, strategies, and operations.

- Level of support for bodies in charge of inspections and internal audits for their catalytic role in fostering an evaluation culture.

- Strategic partnerships between functional entities in charge of evaluations, quality control, and strategies or policies. Such a partnership will translate into a pooling of resources, with the intention of developing capacity-building programs and learning workshops based on outcomes from policy and strategy evaluations.

---

1. OPEV’s Evaluation Community of Practice was launched in 2011 as an internal forum for sharing knowledge. In 2012, the community grew into a forum for sharing and discussing evaluation knowledge with operations personnel working on relevant issues.

Mohamed Manai is Division Manager in the Operations Evaluation Department of the AfDB. Before joining the Bank, Mohamed worked as a Management Consultant for PriceWaterhouse in Kuwait, Tunisia and Morocco. He is trained in finance and holds an MBA degree.
Food for Thought

- How to Speed up Global Development’s Learning Cycle
  *Trevor Davies, KPMG*
- Developing Evaluation Capacity, but Which Capacity?
  *Samer Hachem, AfDB*
- Framework for Understanding NECD Definitions and Terminology
- Checklist for developing a national evaluation system
How to Speed Up Global Development’s Learning Cycle

Trevor Davies leads KPMG’s International Development Assistance Services Global Center of Excellence

I’VE SPENT MORE than 20 years in international development, spending most of that time in fragile or post-conflict states. In that time, I’ve come to see a disconnect between outcomes and intentions.

We all want to ensure the $250 billion committed each year to international development is spent wisely. Yet we do not learn enough, or learn quickly enough, from the work that’s already happening—or happened—on the ground.

This knowledge-management gap means we are constantly missing opportunities to improve the quality of our work and its impact on the world. Several factors contribute to this gap. First, the cycle of evaluation is too slow, and its findings are often out of date or not relevant when they do arrive.

Consider the typical project cycle. Most donors—following the Paris and Busan declarations on aid effectiveness—want to work more closely with the recipient governments and ensure that projects are country-led. At the same time, donors need to ensure these programs meet their own standards and fit within their own agendas.

As a result, the project design phase is an iterative process, and it takes time, in some cases, up to one or two years—or even longer. With the design phase complete, the donor can put in place a baseline, against which progress can be measured. In the meantime, they

… we do not learn enough, or learn quickly enough, from the work that’s already happening—or happened—on the ground… this knowledge management gap means we are constantly missing opportunities to improve the quality of our work and its impact on the world.
A tender for a supplier or implementer, which can take anywhere from three to 12 months, depending on how the contract is structured. Finally, the implementer gets the project underway. Let’s say it continues for five years. When it’s complete, a post-project evaluation is completed.

In total, that’s an eight-year cycle. At what point are you learning lessons and feeding those lessons back into the project? If you begin a project in 2013, you would not learn lessons from it until 2021. So the results are slow to arrive, and when they do, they are considered out of date. A donor will say: “Well, yes, that project was flawed, but we don’t use that framework anymore.”

And this is the second factor contributing to the knowledge-management gap: Our rapid jumping from one development concept to the next.

In the two decades I’ve been involved in development, I’ve seen so many approaches and methodologies. At one time, the new idea was to have stand-alone programs. At another time, it was to have a holistic approach. Then it was spot interventions, then civil society, etc. These ideas were state-of-the-art thinking at the time; guidance notes were sent out, projects designed around them.

In many ways, this rapid evolution in thinking reflects our growing understanding of what works and what doesn’t. But in another sense, we are simply coming up with new ideas to engage constituencies—whether those are taxpayers or lawmakers or check-writing philanthropists. We are finding new ways to tell the same story.

Whatever the cause of this rapid cycling of ideas, the result is that lessons developed over the years through post-project evaluations become much less useful, because they are considered the results of project designs no longer considered valid.

The third factor in our knowledge-management gap is caused by personnel changes at the field level. In many development agencies, people rotate through positions quite rapidly—especially in fragile or post-conflict regions.

Take Afghanistan, where development agencies might post someone for one year. Because that person may take personal leave from the country for at least some of that year, they might be in station only 40 weeks out of the year. Multiply that across an office, and it’s rare that you have a full complement of staff at any one time.

And then, because these are risky environments, you get certain types of people who volunteer for these assignments. Older people with more experience, who often have families, prefer to work at headquarters, so you get younger, less experienced staff. These are complex environments, where the experience and political nuance are needed. A few people are very committed and simply go from country to country, but for many, they go to the field early in their career, and then go back to the head office.

The result is you may have a project conceptualized by one person, designed by a second, commissioned by a third, monitored by a fourth, and evaluated by a fifth. The question is: How can we effectively learn from this project if there is little sense of ownership and continuity?

To bridge this knowledge-management gap, I believe we need to ramp up in-project monitoring
and real time impact evaluation, so that we can learn lessons today and then feed those lessons back into the project.

Of course everyone will say they are doing this type of monitoring, but often it’s just an exercise in ticking a box. It’s not as exciting or sexy as doing the new quasi-scientific things like random control trials, and long-term evaluation—both of which have a role to play—but it does give you the opportunity to course correct.

It’s a bit like the Titanic: Do you want to course correct and miss the iceberg, or do you want to set up a comprehensive review of why it hit the iceberg and what lessons you could learn? Both are valuable approaches but I know which ship I would rather be on.

And with the economic crisis right now putting pressure on aid budgets, we’ve got to make sure that out of every dollar we spend, we are getting the biggest development impact. If something is not working, we need to be able to identify what is not working, take corrective action while the project is ongoing and get back on track.

Some donors are doing in-project evaluation, and doing it well, including the Millennium Challenge Corporation and the U.K.’s Department for International Development.

There’s always a tendency to try and polish things, to get methodologies to 100 percent. In fact, what is important is to learn from what we’re already doing so we can do it better.

KPMG is walking that walk of knowledge management through our new International Development Center of Excellence. Because KPMG is organized around clients, we have teams within KPMG working with key stakeholders in the development community around the world including AusAID, the Asian Development Bank, DfID, SIDA, the United Nations, the World Bank, and many philanthropic organizations.

As you can imagine, this means we may be working on very similar issues, which makes global collaboration and knowledge sharing even more critical as a factor in the success of projects. Through the Center, we aim to bring those insights together, so we can improve development practices.

It’s a small step, and we have a long way to go. But knowledge management is too important not to get right. Because at the end of the day, development is not about coming up with a smart methodology or writing a clever report. It’s about changing people’s lives, they cannot afford to wait for results and neither can the development community.

This column was originally published by Devex Impact. Published here with the author’s permission.

Trevor Davies leads KPMG’s International Development Assistance Services Global Center of Excellence and works across the firm’s global network of member firms. He has more than 23 years of experience working with governments and development agencies.
MY FIRST ENCOUNTER with the concept of capacity development dates a few years back. Coming from the world of management consulting, I used to re-engineer processes, diagnose organizations, and manage change. Moving into the world of development, I was told that our job was to build or develop capacity. But what is this capacity we are supposed to help develop? What kind of assistance can make an institution run itself sustainably? And how am I, a guy who has been doing monitoring and evaluation, supposed to help?

Obviously, I started with definitions. Today it is generally accepted that capacity development has to be seen at three different levels: individual, institutional, and context. While the individual level does not need much explanation, the other two levels are much less obvious.

In a book written in the 1980s, Gareth Morgan proposes eight basic metaphors one might use to look at organizations. These metaphors propose different lenses to analyze and understand how organizations work and evolve. They are machine, organism, brain,
Morgan’s book is based on the premise that almost all our thinking about organizations is based on one or a combination of the eight metaphors. As we look at the institutions and organizations around us in the world of development—development agencies, ministries, civil society organizations, and the like—we can indeed see these metaphors at play.

Think of a machine for example, and you can think structures and processes, and a continuous search for the best output at the lowest cost. Take the organism, and you will see evolving structures, defined by a level of maturity and interactions to function harmoniously. Take the brain, and you will imagine organizations that have mastered knowledge and have learned how to learn.

But Morgan’s eight metaphors also provide a helpful tool to think about capacity. Capacity for a machine-type organization is about executing the same process constantly with efficiency, while capacity for the brain-type organization is about agility and adaptation.

Evaluation, with its dual purpose of accountability and learning, serves the organization as seen through its many facets, the machine but also the brain, the culture, or the political system. Evaluators need to understand these many facets to deliver the right diagnosis, but also, and above all, to select the right path to improve the various capacities of the organization.

To support evaluators, evaluation processes and institutions need to grow and mature over time. Then evaluation can deliver findings and recommendations that will be relevant. What will make a difference is whether this knowledge that evaluation generates helps shape an environment where knowledge is valued as a way of learning and evolving, in other words, where the context of evaluation is conducive to evaluation.

How harmoniously does the evaluation function operate with the other entities? How clear are the processes and structures surrounding evaluation? How well is knowledge packaged and delivered for decision making? How far does leadership accept this knowledge as a guide for action? Beyond individuals and institutions, these are ultimately the key questions evaluation capacity development needs to answer.


Samer Hachem is a manager in the Operations Evaluation Department of the African Development Bank. Prior to joining OPEV, he worked with the AfDB’s Quality Assurance and Results Department, where he focused on monitoring the Bank’s performance on aid effectiveness, and helped introduce the Development Effectiveness Review and a new results management system.

Samer started his career in management consulting, focusing on the financial services sector, and leading monitoring and evaluation in large-scale change management programs. He then joined UNDP HQ in NY, contributing to the roll out of new results-focused management policies and systems. He also worked as a consultant supporting West African countries’ efforts to evaluate their capacities in managing for development results and developing improvement plans.
## Critical success factor checklist for developing a national evaluation (M&E) system

<table>
<thead>
<tr>
<th>Critical success factors</th>
<th>Key considerations for success</th>
</tr>
</thead>
</table>
| Drivers                  | • What is driving the demand for M&E?  
                          | • What are the broad goals of the exercise? |
| Uses                     | • How will M&E information be used, by whom and for what audience(s)?  
                          | • Are there real needs for information that are currently not being met? |
| Leadership                | • Is leadership supportive? Is it leading the way?  
                          | • Is there a ‘champion’ for the M&E exercise? |
| Commitment                | • Is there commitment to launch an M&E exercise and to sustain it?  
                          | • What will serve as the basis to ensure that an M&E system gets implemented and sustained? |
| Resourcing                | • Where will the financial resources to develop systems, and hire and train skilled personnel come from?  
                          | • Will organizations be required to internally reallocate or will they be given new money? |
| Accountability            | • Who will be accountable for ensuring that an M&E system is a functioning part of the public sector?  
                          | • Have roles and responsibilities been firmly established? |
| Technical capacity        | • Is there capacity (data systems and infrastructure) to collect reliable data and report credible information?  
                          | • Is there adequate analytical capacity (skilled personnel)?  
                          | • Are there existing institutions that could serve as credible partners (e.g. national statistical agency, research institutes)? |
| Infrastructure            | • Is there a policy and a set of standards in place that describe roles, responsibilities and expectations for the operation of the M&E system and the use of M&E information?  
                          | • Are the organizations and units that collect and analyse M&E information structured and adequately resourced (budget and human resource capacity)?  
<pre><code>                      | • Is the M&amp;E information that gets reported credible, timely and responding to the priority issues? |
</code></pre>
<table>
<thead>
<tr>
<th>Critical success factors</th>
<th>Key considerations for success</th>
</tr>
</thead>
</table>
| Infrastructure to use M&E information | • Are there formal policies or requirements on how performance monitoring and evaluation information gets used by organizations?  
• What are the incentives within an organization for using M&E information (rewards and/or sanctions)?  
• Are there formal or informal vehicles/mecanisms/forums for reporting, sharing or tabling M&E information?  
• Is results-based performance factored into personnel assessments?  
• Are civil society, the private sector and other social partners actively involved in the national M&E system?                                                                                                                                                                                                                           |
| Oversight                             | • How will the system be monitored over time to ensure that it is functioning as expected and to the level expected?  
• Does the national audit office play any role in monitoring the M&E system and the use of performance information across government?  
• Is there a policy center (e.g. within a central ministry) to monitor implementation?  
• Will the performance of the M&E system be measured and adjusted as necessary?                                                                                                                                                                                                                                                                               |
| Values and Ethics                     | • Is there a formal code of conduct describing accountabilities and expected behaviour for officials in the public sector (e.g. transparency, access to information, fair and balanced reporting, accountability)?  
• Is this code well understood and adhered to by all?  
• Is ‘speaking truth to poverty’ considered appropriate in the public sector and nationally?                                                                                                                                                                                                                                                                         |
| Sustainability                        | • What requirements/safeguards are there to ensure that the M&E system will be made sustainable?                                                                                                                                                                                                                                                                                                                                                                                                 |

About the AfDB: The overarching objective of the African Development Bank Group is to foster sustainable economic development and social progress in its regional member countries (RMCs), thus contributing to poverty reduction. The Bank Group achieves this objective by mobilizing and allocating resources for investment in RMCs and providing policy advice and technical assistance to support development efforts.

The mission of the Operations Evaluation Department is to enhance the development effectiveness of the AfDB in its regional member countries through independent and instrumental evaluations and partnerships for sharing knowledge.

Operations Evaluation Department, African Development Bank
Website: http://operationsevaluation.afdb.org/
Write to us: opevhelpdesk@afdb.org

From Experience to Knowledge …
From Knowledge to Action … From Action to Impact