Capacities and Capabilities

Chapter 4

Incorporating the Sustainable Development Goals in National Evaluation Capacity Development

Indran Naidoo and Ana Rosa Soares

Abstract. *This chapter discusses the efforts of the United Nations Development Programme to develop national evaluation capacities through the biannual conferences and actions promoted by its Independent Evaluation Office. The paper also looks into lessons learned from implementing the Millennium Development Goals that could be useful in evaluating the Sustainable Development Goals (SDGs). It further outlines directions and priorities for incorporating the SDGs in national evaluation capacity development efforts, building on what emerged from the consultations that took place during the Fourth International Conference on National Evaluation Capacities in Bangkok, which was jointly organized with the 2015 Global Assembly of the International Development Evaluation Association.*

At a United Nations (UN) summit in September 2015, member states adopted the 2030 Agenda for Sustainable Development, "a plan of action for people, planet and prosperity" that "seeks to strengthen universal

Indran Naidoo, United Nations Development Programme, indran.naidoo@undp.org;
Ana Rosa Soares, United Nations Development Programme, ana.soares@undp.org.

peace in larger freedom" (UN 2015c). The 2030 Agenda commits all countries and various stakeholders to work together to "free the human race from the tyranny of poverty and want and to heal and secure our planet," to address inequality and injustice, and to ensure "that no one will be left behind." The 2030 Agenda presents an integrated plan of action with a vision and principles for transforming our world as set out in the results framework of 17 Sustainable Development Goals (SDGs) and 169 targets, with quantitative and qualitative objectives for the next 15 years; a means of implementation and global partnership; and a follow-up and review process. The follow-up and review framework calls for accountability to the people, national ownership, and country-led evaluative processes. Evaluation practice will provide an important means for raising the voice of stakeholders in this process to inform, support, measure, and assess whether development progress around the SDGs is relevant, sustainable, and equitable. Developing national evaluation capacities will be necessary in order to ensure that the follow-up and review process adds value to the implementation of the SDGs.

UNDP DEVELOPING NATIONAL EVALUATION CAPACITIES TO EVALUATE DEVELOPMENT GOALS

In 2015, the Independent Evaluation Office (IEO) of the United Nations Development Programme (UNDP) and the Royal Thai Government cohosted the Fourth International Conference on National Evaluation Capacities (NEC) in Bangkok, in collaboration with the UNDP Regional Bureau for Asia and the Pacific. The conference was organized jointly with the 2015 Global Assembly of the International Development Evaluation Association (IDEAS).

This conference was the fourth in a series of NEC conferences that have recognized UNDP for its distinct focus on supporting the governments with which UNDP works across the globe, as part of an IEO strategy to support the development of national evaluation capacities. The NEC conferences are held by UNDP every two years, each time in a different region, in partnership with a host government.

The model of the NEC conference has evolved over the years, drawing on lessons learned and emerging demands: but of essence is the focus on supporting governments to build their accountability capacities, of which evaluation is a key part. Over the last 10 years, the event has involved different partners, each of which has provided a particular emphasis and served to enrich the discussions, making it a key global evaluation event.

Each time around, support is focused on a specific region and uses different formats of exchange to promote commitment, cooperation, and action between and beyond senior government officials and to encourage the sharing of responsibility with other key players in the evaluation community. Much effort has also been invested in promoting continued engagement with past participants and institutions to deepen dialogues and foster continuity, partnership, and learning.

The NEC conferences are part of a broader architecture, in which the United Nations Evaluation Group (UNEG) plays a role by bringing together UN agencies and development partners to collaborate with each other. These

occasions serve to enhance the understanding and appreciation of evaluation as a powerful tool of public accountability and learning. They also help to advance the evaluation discourse globally and to align it with a strong call for cooperation in "building capacity for the evaluation of development activities at the country level" highlighted in the UN General Assembly Resolution 69/237 (UN 2015a). This resolution invites the entities of the UN development system, with the collaboration of national and international stakeholders, to support, upon request, efforts to further strengthen the capacity of member states for evaluation, in accordance with their national policies and priorities. Through UNEG, the UN promotes professional norms and standards for evaluation. In addition, UN entities and partners use evaluation to support accountability and program learning; to inform UN systemwide initiatives and emerging demands; and to benefit from and contribute to an enhanced global evaluation profession. The UN plays a particularly important role in enhancing national capacities to monitor and evaluate progress in poverty eradication and other internationally agreed-upon development goals, and therefore its key responsibility in supporting the SDGs.

UNDP's value added in evaluation has been its contribution as the secretariat and cochair of UNEG, but above all the strong independent mandate of its IEO to evaluate its contributions to development. IEO works with UNDP country offices and bureaus for national evaluation capacity development with a clear division of roles and responsibilities. IEO is responsible for developing guides and standards, and for promoting national evaluation capacity development through discussion forums such as the NEC conferences and associate initiatives of knowledge exchange. UNDP program units are further responsible for following up on the outcomes and partnerships of the conferences, and for supporting more specific programs and partnerships to develop national evaluation capacities in the medium to long term. This division of roles and responsibilities ensures IEO's independence and ability to credibly evaluate the results of UNDP's contributions.

UNDP believes that when appropriately tailored to national circumstances and priorities, the evaluation function can be an effective country-led vehicle for greater citizen accountability that can accelerate progress toward national SDG priorities, drawing on contributions from indigenous peoples, civil society, the private sector, and other stakeholders, including national parliamentarians (UNDP 2016a). Governments are engaged by UNDP to identify national evaluation partners, especially during the NEC conferences, but also as partners in certain evaluations. In these processes they develop their national evaluation capacities to promote greater accountability, learning, and development effectiveness in their countries.

UNDP has been supporting a range of activities to promote national evaluation capacity development, based on the UNDP definition of capacity development as an endogenous process through which individuals, organizations, and societies obtain, strengthen, and maintain the capabilities to set and achieve their own development objectives over time. Such a process can be described as country-owned if it is operated in a dynamic change process with reflection and learning, and if it is gradual, opportunistic, and adaptive to varying circumstances.

Using this frame of reference, the IEO NEC strategy approaches evaluation capacity development and the purpose and meaning of evaluation from a country, as opposed to a donor, perspective. In this regard, the purpose of evaluation goes beyond accountability to donors, to encompass public sector efficiency and accountability to the citizens of the country. The purpose of evaluation embraces other significant institutional and national goals for learning, and the development of innovation and social capital, knowledge assets, and the intellectual capital needed for growth, development, and contribution to global advancement.

UNDP has been successful in linking theory with practice, vision and ideals with realities, and in the methods we have sought to do the bridging. The Fourth NEC Conference provides a clear example of this approach with the theme "Blending Evaluation Principles with Development Practices to Change People's Lives." Together with IDEAS, this conference was an important opportunity to engage decision makers, academics, practitioners, and the UN community in global dialogue and advocacy around evaluation and the SDGs. More than 450 participants from 100 countries and from three key evaluation networks—UNEG, the Evaluation Cooperation Group of the Multilateral Development Banks, and the Evaluation Network of the Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD DAC)—participated in the conference, indicating the important role evaluation should play in shaping and contributing to the SDGs during the upcoming 15 years. The conferences are also an important opportunity for countries interested in South-South and South-North cooperation to find solutions together for challenges that have no ready-made answers. All can learn from previous experiences, such as what was advanced during the decade of the Millennium Development Goals (MDGs), that can be useful for the SDGs.

In Brazil in 2013, in a previous iteration of the conference, participants discussed solutions to challenges related to the independence, credibility, and use of evaluations. The conference produced 18 NEC commitments to further enhance national evaluation capacities, and encouraged creating greater accountability by setting goals for each country's NEC journey.[1] The 18 NEC commitments centered around four main strategies to build national evaluation capacities:

- Promote evaluation use through in-country and global advocacy
- Define and strengthen evaluation processes and methods
- Engage existing and new stakeholders in exchange and collaboration
- Explore options for different institutional structures for managing evaluations

In 2015, IEO published a baseline assessment of the countries that have participated in the NEC conferences series in order to document where each

[1] The NEC commitments can be found at http://www.nec2013.org/.

country stood, and in what direction they were moving regarding national evaluation capacities (IEO UNDP 2015c). This assessment found a variety of institutional settings and legal frameworks among the countries, reflecting a variety of government interests, political contexts, and national developmental stages. These granular aspects of national evaluation capacities are complex, and intrinsically linked to each country's development agenda: therefore they need to be taken into consideration and incorporated into the development of future evaluation agendas.

Over the years, one key lesson was learned through the NEC conferences and the process of promoting and implementing NEC commitments: without clear goals and the appropriate follow-up, governments and partners have a hard time focusing their attention on, and committing to the changes and long-term investments needed in order to build national evaluation capacities. The 18 NEC commitments served as a conversation starter for NEC participants to go back to their countries and reconsider their key national evaluation capacities and needs. These commitments were not necessarily the 18 initial commitments of the 2013 NEC conference, but included specific commitments that the countries agreed made more sense for their unique national contexts, and were therefore equally important.

Expanding on the 18 NEC commitments, in a global partnership effort for 2015, the International Year of Evaluation (EvalYear), the NEC 2015 conference focused on gathering information and commitments from participants to develop a new set of NEC commitments. The outcome was the Bangkok Declaration, a much expanded format that went beyond NEC and incorporated elements that also focused on the evaluation profession and global issues.[2] The declaration later contributed to another relevant document, the Global Evaluation Agenda, which was the first ever long-term global vision for evaluation.[3]

The Bangkok Declaration was a collective statement of all participants of the joint 2015 NEC conference and IDEAS Global Assembly: it is an expression of aspirations grounded in the community of practice of professional development evaluation. It is not legally binding on individuals or governments, but it seeks to capture key principles, give a sense of common purpose and understanding, and frame a vision of joint action in future support of individual, professional, and national evaluation capacity as countries shape their responses to the 2030 Agenda.

CONVERGENCE OF THE MDGS AND THE SUSTAINABLE DEVELOPMENT TRACKS FOR THE SDGS

Recognizing the intrinsic linkage between poverty eradication and sustainable development, during the General Assembly Special Event in September 2013, UN member states requested that the Open Working Group and the

[2] The Bangkok Declaration can be found at http://web.undp.org/evaluation/nec/nec-2015_declaration.shtml.

[3] The Global Evaluation Agenda 2016–2020 can be found at http://www.evalpartners.org/global-evaluation-agenda.

Committee of Experts on Sustainable Development Financing produce inputs for the post-2015 negotiations of the SDGs.

In August 2014, the Open Working Group submitted its proposal for a set of 17 SDGs, along with 169 associated targets. At the same time, the Committee of Financing Experts produced a set of recommendations on sustainable development financing. In December 2014, the Secretary-General submitted to UN member states his synthesis report, combining the intergovernmental proposals and the full range of inputs from both tracks.

UN member states agreed that the proposed SDGs would form the basis for intergovernmental negotiations of the post-2015 agenda with a text of the new agenda entitled "Transforming Our World: The 2030 Agenda for Sustainable Development," for adoption by the Post-2015 Summit held in New York September 25–27, 2015. The text included a declaration, 17 SDGs and 169 targets, and components on the means of implementation, the global partnership, and a follow-up and review process. The 2030 Agenda was structured around five "Ps"—people, planet, prosperity, peace, and partnership—and its set of 17 SDGs were officially adopted by the summit on September 25. The comprehensive nature of this new agenda has effectively reaffirmed this convergence, aligning the processes and the scope, and leading to a holistic approach to development.

Three other complementary processes ran alongside the Post-2015 process: one with a focus on disaster risk reduction, another on financing for development, and a third focused on climate change. The 2030 Agenda became an umbrella agreement for these other three agreements as well.

In March 2015, UN member states adopted the Sendai Framework for Disaster Risk Reduction (2015–2030) during the Third World Conference on Disaster Risk Reduction held in Japan. This framework, which is the result of several years of consultations and several months of intergovernmental negotiations, contains seven targets and four priorities for action to reduce negative impact, build resilience, and strengthen related international cooperation. The 2030 Agenda explicitly recognizes the importance of disaster risk reduction, and makes reference to the Sendai Framework and the need for development of holistic disaster risk management at all levels in its Goal 11.

In July 2015 in Addis Ababa, UN member states held the Third International Conference on Financing for Development, organized as a follow-up to the Monterrey Consensus and Doha Declaration. This conference led to an agreement entitled "Addis Ababa Action Agenda," which identifies the key action areas needed in order to provide the means and create an enabling environment for implementing the SDGs. The text of the adopted 2030 Agenda recognizes the concrete policies and actions agreed to in Addis Ababa as supporting, complementing, and contextualizing the means of implementation targets of the SDGs, and the Addis Ababa Action Agenda itself as an integral part of the 2030 Agenda for Sustainable Development.

The third complementary process is the UN Framework Convention on Climate Change (UNFCCC). The 21st Conference of the Parties (COP21) that was held in Paris in late 2015 featured negotiations toward the first universal, legally binding global agreement on climate change, now known as the Paris Agreement. The Paris Agreement, which is due to enter into force in 2020,

contains an action plan that could allow UN member states to limit global warming to well below 2°C, and aims to limit it to 1.5°C. In addition to actions aimed at the reduction of emissions, it also covers issues related to adaptation, support, loss and damage, and transparency and stocktaking. Such an agreement is explicitly mentioned in the proposed SDGs, and the UNFCCC is acknowledged as the primary forum for these negotiations.

APPLYING LESSONS LEARNED FROM THE MDGS TO THE SDGS

The post-2015 negotiations, and the work to design the SDGs, were led by member states from the outset. The outcome is the result of a consistent global participatory process in which representatives from countries, academia, civil society, and the private sector together formulated the SDGs. The SDGs contain much that critics said was missing in the MDGs.

One thing that was learned from the experience of the MDGs was the importance of reporting and monitoring. However, the 2030 SDG Agenda has a much wider scope than the largely social goals of the MDGs, and takes into greater consideration the need for economic, social, and environmental sustainability. It also recognizes the importance of peaceful societies. With the MDGs, the question was: What are the goals that are lagging the most, where are the gaps, and how can we fill them? With the SDGs, given the fact that the breadth of the 2030 Agenda implies a need to break down silos and adopt an integrated approach to development interventions, the question has become more evaluative. Reporting and monitoring are insufficient; countries need evaluations in order to answer this question: What are the actions required to accelerate progress across a broader range of interlinked goals? Addressing this question requires thinking through the connections and synergies across the goals, and pointing out how actions in one area affect other areas. Evaluative tools are also required to assess and manage trade-offs, and in this context "evaluation methods will need to determine whether the right choices were made to achieve possibly conflicting desirable outcomes, and how the different outcomes should be valued" (Heider 2015).

While much has been achieved during the MDG implementation period, a key criticism of the MDGs was that there was insufficient attention paid to generating evidence on achievements and particularly learning from challenges.[4] Much greater focus has been on monitoring and reporting, with many countries publishing national and also subnational MDG progress reports, while evaluation of which policies and interventions have worked and which have not were often only conducted at a later stage, and as part of designing MDG acceleration frameworks. MDG progress was largely tracked at the aggregate level, masking disparities in performance and disguising rising inequalities. In order to move forward in such a way as to ensure that no one is left behind, a better understanding of why and how certain policy choices and interventions affect different segments of society will be imperative. Recognizing that "only

[4] See EvalSDGs, http://www.unicef.org/evaluation/files/EvalSDG_Overview_Paper_8-12-15_1-pager.pdf.

by counting the uncounted can we reach the unreached" (UN 2015b), SDG targets should be met for all nations, peoples, and segments of society, and should "reach the furthest behind first," but only by evaluating trends, and contributing and hindering factors, can we assess whether progress has been or can be made relevant, sustainable, and equitable.

The 2030 Agenda is also a much more ambitious agenda than the MDGs, aspiring toward the goals of the elimination of poverty and universal access to benefits: this requires addressing the root causes of exclusion and deprivation, which are often deeply embedded in economic, social, and political marginalization. Another key lesson of the MDG implementation was that early strategic planning is important in laying the groundwork for long-term progress, because putting into place priority actions at an early stage can have multiplier effects on development outcomes (IEO UNDP 2015b). Targets associated with the MDGs were only shaped over time, while financing the MDGs was discussed in Monterrey two years after the Millennium Declaration. In contrast, the inclusion from the outset of a detailed results framework in the 2030 Agenda presents an opportunity for early action to link results and resources for results-based management.

From the beginning, the follow-up and review mechanism of the SDGs will also allow for early adjustments, course corrections, and enhanced results. In addition, the "MDG monitoring experience has clearly demonstrated that effective use of data can help galvanize development efforts, implement successful targeted interventions, track performance and improve accountability" (UN 2015b, 10). The MDG framework also strengthened the use of robust and reliable data for evidence-based decision making, with many countries integrating the MDGs into their national priorities and development strategies. Country ownership, leadership, and the participation of a wide range of stakeholders have been vital to ensure MDG progress and accountability.

PRIORITIES FOR NATIONAL EVALUATION CAPACITY DEVELOPMENT IN THE POST-2015 SDG DEVELOPMENT AGENDA

The 2030 Agenda states that "Governments have the primary responsibility for review, at the national, regional and global levels, in relation to progress made in meeting the goals and targets over the next fifteen years" (UN 2015c). With the explicit follow-up and review mechanism of the SDGs, countries will need to go beyond the usual monitoring and tracking of the MDGs, and tackle evaluations.[5]

Given the complexity of the SDGs—17 goals, 169 targets, and 230 indicators—the evaluation community has to be prepared to support an SDG platform for measurement, and for improving national evaluation capacities to contribute to accountability and learning. In addition, investment in qualitative

[5] The universal nature of the 2030 Agenda and the changing dynamics of development finance and development cooperation also present an opportunity to move from donor-driven to country-led evaluation.

assessment and careful design of national and international platforms and networks for dialogue, information sharing, and debate, with particular attention given to evidence provided by diverse domestic actors, may become central to achieving the SDGs.

The challenge of implementation points to the need for learning what works and what does not; which factors influence and hamper success; which aspects can risk sustainability, under which contexts; and how to break down silos and promote an integrated approach in order to achieve the most effective and efficient results. With that in mind, four overarching priorities emerged from the consultations that took place during the last NEC conference in Bangkok, building on the discussions from previous NEC conferences about independence, credibility, and the use of evaluations.

Promoting country-owned, country-led evaluations, with an emphasis on their use in influencing policies. One important priority is to respond to national circumstances, to support existing national systems and to avoid duplication of efforts and the famous "reinventing the wheel." Doing this entails a shift from donor-driven evaluations to country-owned evaluations and developing national evaluation capacities. This process should not be donor-driven but rather localized, contextualized, and culturally sensitized.

Critical for national ownership of evaluations is the need to raise the demand for evaluations, and not just focus on supply. A successful use of evaluations to inform policy, and to promote a change in mindsets within organizations and governments, can be used to advocate for a prominent role for evaluation in the implementation of the SDGs, for learning, and ultimately, to bring about improvement in people's lives.

Developing and strengthening evaluation process and methods. A second priority identified at the conference is developing new methods for evaluating progress toward, and the impact of, the SDGs. The 2030 Agenda is committed to developing broader measures of progress to complement gross domestic product. But how do we measure sustainability with the SDGs in mind? Environmental protection is only one of the means to achieve continued ecosystem services to mankind of clean air, water, healthy food, and freedom from disease. Sustainability requires an adaptive dynamic balance between the social, economic, and environmental domains. The SDGs require seeing economic growth, social inclusion, and environmental protection as mutually reinforcing. Whether our measurement and evaluation tools are sufficiently sophisticated to provide evidence on whether such a dynamic balance has been reached, or is within reach, and whether it is adaptive enough to change when necessary are great challenges. Methods that capture social inclusion and environmental protection need to be found in order to assess and evaluate sustainability.[6]

[6] There has been increasing interest from governments around the world in using innovative techniques to get better feedback from citizens on the effectiveness of their policies and programs, and to improve equity, sustainability, and accountability.

Engaging existing and new stakeholders in exchange and collaboration. A third priority identified was the importance of promoting more diverse partnerships and greater cooperation between governments, civil society, parliaments, and the private sector, in order to increase the awareness and use of evaluations. Traditional North-South aid models are playing a increasingly small role as private sector and national government resource flows increase, and the evaluation community advocates for more country-driven evaluations. There is a need for more dialogue in order to improve cooperation between the public and private sectors, to create networks and platforms for information and knowledge sharing, and to involve representatives of the private sector, parliamentarians, policymakers, legislators, and individual citizens.

The conference also stressed the importance of citizens as stakeholders, and the importance of raising awareness among citizens of the SDGs and the role of evaluation. There is growing awareness of the importance of people's engagement in monitoring and evaluation, and in accountability mechanisms.

Institutional structures for the evaluation of the SDGs. The NEC Conference revealed that we still have a long way to go in understanding how we integrate the evaluation of the SDGs into institutional structures. Almost every SDG is covered by national policy, so the question governments now face is how to monitor and evaluate all these policies and SDGs without duplicating and wasting resources. The holistic and integrative nature of the SDGs is not reflected in its structure and division into 17 goals. Governments may be tempted to divide out responsibility for the SDGs to respective line ministries, and the integrative perspective may be lost as a result. We need to sustain the discourse on the need to work on all SDGs also in evaluation. In an environment where the resources needed to deliver on the SDGs are scarce, evaluation will continue to gain prominence as a means of ensuring accountability for the use of those resources, and can help nations learn what works best under which contexts, in order to ensure effectiveness. The right institutional structures and national evaluation capacities will be key to the success of these processes.

CONCLUSION

The 2030 Agenda spells out the ideals and goals that will require evaluators and the development community to engage with in addressing a variety of interrelated, complex, and challenging issues, and to be competent at multiple levels in order to make significant contributions. If evaluators are to help give voice to people and countries in a global context where inequality persists at multiple levels, there is a need to start thinking about evaluation of the SDGs now, rather than as an afterthought. The SDGs contain a vision that combines a human capability approach to development with modern

UNDP, through its Innovation Fund, has implemented a number of prototypes with partners that harness technology to improve sustainability and accountability.

reconstructions of traditional economic models of growth. In responding to the SDGs focus on inequity, and in service to the principle of "no one left behind," the evaluation function can bring methodological validity as well as the legitimacy to empower people as effective evaluation processes help promote social action for development.

Achieving the SDGs depends on country-led evaluations that will produce evidence of whether the outcomes and impacts of policies, programs, and projects are equitable, relevant, and sustainable. Such evidence is useful not only in demonstrating public sector accountability, but also in focusing the attention of civil society and governments on enhancing learning, adaptive management, and innovation. Evaluation does not only identify "what works and what doesn't," or simply answer the question of whether we did or did we not achieve our objectives. Its real value is that it can be coupled to learning. For that, the learning and knowledge highlighted in evaluations needs to be used beyond simple accountability for strategic planning and adaptive management. Evaluation is a dynamic and ongoing process that continues to evolve, and is vital to support improving efforts, results, and development.

Therefore, supporting national evaluation capacity development is key to enabling mutual accountability among countries, and promoting learning to further the effective achievement of the SDGs, while ensuring that no one is left behind.

IEO is proud of the role UNDP has played in supporting development, and stands ready to work with partners to advance in supporting the development of national capacities for evaluating progress toward the SDGs.

In 2017, the Fifth International Conference on National Evaluation Capacities will take place. The evaluation community should be intensely engaged in discussing how to assess the equitability of SDG outcomes for marginalized populations; how to measure and evaluate new themes that are integral to the SDGs; and how to assess the effectiveness of integrated approaches, in order to understand what works best and under which contexts, to expedite progress toward and the achievement of the SDGs.

BIBLIOGRAPHY

Bester, Angela. 2015. "Capacity Development." Report prepared for the United Nations Department of Economic and Social Affairs 2016 Quadrennial Comprehensive Policy Review.

BetterEvaluation. n.d. "Steps in Planning and Managing an Evaluation."

Goldman, I., J.E. Mathe, C. Jacob, A. Hercules, M. Amisi, T. Buthelezi, et al. 2015. "Developing South Africa's National Evaluation Policy and System: First Lessons Learned." *African Evaluation Journal* 3 (1). http://dx.doi.org/10.4102/aej.v3i1.107.

Hauge, A. 2003. "The Development of Monitoring and Evaluation Capacities to Improve Government Performance in Uganda." Operations Evaluation Department ECD Working Paper No. 10, World Bank, Washington, D.C.

Heider, Caroline. 2015. "Evaluation beyond 2015: Implications of the SDGs for Evaluation." Blog post. Independent Evaluation Group, World Bank, September 15.

IEO UNDP (Independent Evaluation Office of the United Nations Development Programme). 2015a. "Blending Evaluation Principles with Development Practices to

Change People's Lives." Proceedings from the Fourth International Conference on National Evaluation Capacities. UNDP, New York.

———. 2015b. *Evaluation of the Role of UNDP in Supporting National Achievement of the Millennium Development Goals*. New York: UNDP.

———. 2015c. *Towards a Baseline Study: Insights on National Evaluation Capacities in 43 Countries*. New York: UNDP.

Mackay, K. 2007. *How to Build M&E Systems to Support Better Government*. Washington, DC: Independent Evaluation Group of the World Bank.

Naidoo, I. 2014. "The Contribution of Country Level Evaluation in Strengthening UNDP's Role in State-Building." African Development Bank Group, Abidjan.

Nilsson, Måns, Dave Griggs, Martin Visbeck, and Claudia Ringler. 2016. "A Draft Framework for Understanding SDG Interactions." International Council for Science.

OECD DAC (Organisation for Economic Development Development Assistance Committee). 1991. "DAC Principles for the Evaluation of Development Assistance." OECD, Paris.

Ofir, Zenda, Thomas Schwandt, Stefano D'Errico, Kassem El-Saddik, and Dorothy Lucks. 2016. "Five Considerations for National Evaluation Agendas Informed by the SDGs." IIED Briefing, International Institute for Environment and Development, London.

Schwandt, Thomas, Zenda Ofir, Dorothy Lucks, Kassem El-Saddick, and Stefano D'Errico. 2016. "Evaluation: A Crucial Ingredient for SDG Success." IIED Briefing, International Institute for Environment and Development, London.

Schwandt, Thomas, Zenda Ofir, Stefano D'Errico, Kassem El-Saddik, and Dorothy Lucks. 2016. "Realizing the SDGs by Reflecting on the Way(s) We Reason, Plan and Act: The Importance of Evaluative Thinking." IIED Briefing, International Institute for Environment and Development, London.

UN (United Nations). 2015a. "Building Capacity for the Evaluation of Development Activities at the Country Level." A/RES/69/237. UN General Assembly, New York.

———. 2015b. *The Millennium Development Goals Report 2015*. New York: UN.

———. 2015c. "Transforming Our World: The 2030 Agenda for Sustainable Development." A/RES/70/1. UN General Assembly, New York.

———. 2016a. "Interim United Nations Development Assistance Framework Guidance." UN, New York.

———. 2016b. "Ministerial Declaration of the High-Level Segment of the 2016 Session of the Economic and Social Council on the Annual Theme "Implementing the P-2015 Development Agenda: Moving from Commitments to Results." UN, New York.

———. 2016c. "Progress towards the Sustainable Development Goals." Report of the Secretary-General. UN, New York.

———. 2016d. "Report of Executive Summaries of Volunteer Countries on Voluntary National Reviews for the High-Level Political Forum." UN, New York.

———. 2016e. "Report of the High-Level Political Forum on Sustainable Development Convened under the Auspices of the Economic and Social Council at Its 2016 Session." UN, New York.

UNDP (United Nations Development Programme). 2009. *Capacity Development: A UNDP Primer*. New York: UNDP.

———. 2013. *Solutions Related to Challenges of Independence, Credibility and Use of Evaluation: Proceedings from the Third International Conference on National Evaluation Capacities*. New York: UNDP.

———. 2016a. "The UNDP Evaluation Policy." DP/2016/23. United Nations, New York.

———. 2016b. "UNDP Support to the Implementation of the 2030 Agenda for Sustainable Development." Policy and Programme Brief, UNDP, New York.

UNEG (United Nations Evaluation Group). 2012. "National Evaluation Capacity Development: Practical Tips on How to Strengthen National Evaluation Systems." UNEG, New York.

———. 2016. "Norms and Standards for Evaluation." UNEG, New York.

World Bank. 1999. *Evaluation Capacity Development: A Diagnostic Guide and Action Framework*. Washington, DC: World Bank.

———. 2012. "Guide to Evaluating Capacity Development Results." A collection of guidance notes to help development practitioners and evaluators assess capacity development efforts. World Bank, Washington, DC.

Chapter 5

Professionalizing Evaluation – A Golden Opportunity

Linda Morra Imas

Abstract. *This chapter considers the strong mandate for evaluation provided by the 2030 Agenda for Sustainable Development; and there are laudable efforts to strengthen national evaluation capacities in line with the mandate. However, there is still a lack of clarity on what such capacities look like in the context of the Sustainable Development Goals (SDGs), and what this means for evaluator competencies. Evaluators can view this situation as an opportunity to move forward on establishing core competencies for evaluators that reflect the SDGs, as well as to develop a clearer vision of national capacities to evaluate them.*

The good news is that the field of evaluation has a strong mandate, and is responding to it: on January 1, 2016, the 17 Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development—which were adopted by world leaders in September 2015 at a historic UN Summit—officially came into force (UN 2015). Over the next 15 years, these new goals, which universally apply to all countries, will mobilize efforts to end all forms

Linda Morra Imas, International Program for Development Evaluation Training, lindag1000@gmail.com.

of poverty, fight inequalities, and tackle climate change, while ensuring that "no one is left behind."

Ensuring that no one is left behind means strengthening the voices and power of the most marginalized members of society—the disabled, the young, women, the poor—and challenging some of the most vested interests, such as those of energy producers. These are no small tasks, and there are some who believe the goals will not be realized by 2030.[1] But even to know how we are progressing toward these goals, evaluation of poverty and inequity, both within and across countries, is clearly needed.

The SDGs also provide an international mandate for evaluation, and continue to propel an evaluation capacity-strengthening movement that began with the Millennium Development Goals (MDGs). From these and from other ongoing influences, evaluation societies, known as voluntary organizations for professional evaluation (VOPEs) have sprung up at all levels—national, subnational, regional, and international—from just 15 in 1999 to 151 verified VOPEs by 2016.[2] Some have developed, or are in the process of developing, competency systems for their memberships, and a few have moved to credentialing or qualification systems. Both are ways of attempting to increase the professional status of evaluation, among other things.

But is this too little, and is it moving too slowly? The SDGs are multidimensional and tend to require mixed methods and complex, system-level insights. Without agreement on basic core competencies for the profession, and by letting "a thousand flowers bloom," are we missing a golden opportunity to advance the professionalization of evaluation? This chapter explores these issues.

A STRONG INTERNATIONAL MANDATE FOR EVALUATION

It cannot be overemphasized that the 17 SDGs, with their 169 targets, each with multiple indicators, are a first step in requiring all countries—not just the so-called developing countries—to set their own national agendas and strategies in collaboration with stakeholders. In this context, *all* countries are "developing" countries, facing common issues. The 2030 Agenda for Sustainable Development calls for follow-up and review processes that examine progress toward achieving the SDGs at the country, regional, and international levels. Follow-up and review processes are to be "rigorous and based on evidence, informed by country-led evaluations and data which is high-quality, accessible, timely, reliable and disaggregated by income, sex,

[1] See, e.g., Berliner (2015). Berliner and his team selected one key target for each of the 17 goals. Using projections from leading organizations, they predicted that not a single goal will be reached by 2030 if current trends continue.

[2] The International Organization for Cooperation in Evaluation (IOCE), the umbrella organization for evaluation organizations, lists the number of VOPEs as 188 in 2013 on its home page (ioce.net). Creating VOPEs may be easier than sustaining them. IOCE maintains an Excel database on VOPEs and, as of May 7, 2016, reports 151 verified VOPEs. This still represents huge growth from 1999.

age, race, ethnicity, migration status, disability and geographic location and other characteristics relevant in national contexts" (UN 2015). As indicated by EvalSDGs, a network of policy makers, institutions, and practitioners who advocate for evaluation of the SDGs, the initial focus has been necessarily on how to define and measure progress using indicators. However, it is acknowledged that measurement is not enough: "monitoring must be accompanied by evaluation that addresses the complexity of the SDGs and how they are achieved" (Schwandt et al. 2016).

WHAT DOES EVALUATION OF THE SDGS ENTAIL?

How to evaluate the SDGs is not so clear when it comes to the specifics, and relatively few have tried to articulate a clearer vision. Taking the lead, the International Institute for Environment and Development (IIED), in partnership with the network EvalSDGs, has been producing a series of briefs on this topic. At a basic level, the April 2016 brief indicates that evaluation uses monitoring data, but adds that it is "primarily concerned with how well implementation, outputs, and development outcomes were achieved, as well as with determining long-term development impact." It also says that "Evaluation asks why targets were achieved or not achieved and what can be done to improve the likely success of future initiatives" (Schwandt et al. 2016a, 2). But these generic statements about evaluation are followed with an emphasis on the interconnectivity of the SDGs that leads to the need to think about the evaluation of complex systems, rather than the evaluation of a single policy, program, or project.

More specifically, in terms of the SDGs, the brief notes that because the SDGs are interrelated in such complex ways, they present "wicked" problems for evaluation. For example, the aim of reducing income inequality (SDG 10) cannot be neatly separated from the aim of ensuring healthy lives and well-being (SDG 3). Initiatives to address such problems are themselves complex. They may involve "long causal chains with many intermediate outcomes, or outcomes that can only be understood using a 'causal package' approach that examines contributions from multiple interventions, contexts, or agencies..." (Schwandt et al. 2016a, 3). The implication for evaluation is that skills in new evaluation methodologies that draw on systems thinking may be needed. Another brief focuses on critical thinking skills as essential for conducting evaluations that analyze arguments, weigh evidence, and assess claims (Schwandt et al. 2016b). Being able to conduct country-led evaluations that assess sectoral, thematic, and holistic national policies, and that reflect whether a problem was correctly identified, the intended effects achieved, and whether unintended effects—either positive or negative—occurred is another part of the skills picture. Also, one must determine that outcomes and impacts are equitable, relevant, and sustainable. Because evaluations are not only at the national level but country led, stress is additionally placed on partnerships and evaluation capacity building.

Others—for example, Patton's "blue marble evaluation"—have put an emphasis on the need for evaluation from a global perspective, and the ability to evaluate adherence to principals such as human rights, gender equity,

inclusiveness, and sustainability, as well as the importance of maintaining the independence of evaluations (UNEG 2016c).[3] Yet others focus on the evaluation process itself, stressing the involvement of stakeholders, and qualities such as mindfulness, inclusiveness, and facilitation skills (Catsambas 2016).

Those who favor experimental and quasi-experimental approaches to evaluation have raised their voices to stress that information must be "evidence-based" and "rigorous." Some, such as the organization 3ie (International Initiative for Impact Evaluation), are promoting the use of findings from systematic reviews as a sounder base than individual studies for policies and programming (White 2015). This is because many lessons come from evaluation anecdotes, correlational data, and strong counterfactual data. These lessons need to be sorted out in order to use the most rigorous findings, and to see how the effects in one setting hold in other settings and over time.

From these writings, we could draw up an impressive and long list of skills that might be needed by those seeking to evaluate SDGs. But we also know that countries currently have not only widely different levels of evaluation capacity, but also wide variability in the availability of evaluation training. Variability in evaluation and in the quality of evaluation likely will be the main stories in efforts to evaluate the SDGs today, such that it might be difficult to have a coherent picture beyond indicator data. The ability to look at and address SDG issues from a global, regional, or even subnational perspective is limited. Professionalization of evaluation could provide an opportunity for a more level playing field by identifying global core competencies, and focusing training on building those competencies.

THE NEED FOR NATIONAL EVALUATION CAPACITY STRENGTHENING

Even at a basic level, it has been apparent for many years that many countries need assistance in developing their national evaluation capacities. The MDGs had building national evaluation capacity as an emphasis, and that emphasis is continued with the SDGs. For example, the World Bank, through its Independent Evaluation Group, fostered the development and implementation of the long-running International Program for Development Evaluation Training (IPDET), and it also partners in the Regional Centers for Learning on Evaluation and Results (CLEAR). With the SDGs, UN organizations are making increased efforts to support VOPEs. To illustrate, the International Organization for Cooperation in Evaluation (IOCE) has as its mission to support VOPEs in contributing to good governance, effective decision making, and strengthening the role of civil society. Under an agreement with UNICEF, IOCE

[3] At a lunch presentation at IPDET in June 2015, Patton floated the idea of "blue marble evaluators." He was received enthusiastically and based on the reaction developed a proposal to move forward the perspective of a complex, dynamic, and interconnected world system. The "blue marble" perspective means thinking globally, holistically, and systematically. Evaluators need special perspectives and competencies to engage and evaluate these global change efforts.

launched the EvalPartners Peer-to-Peer Program, which encouraged two or more VOPEs to form partnerships to strengthen their capacities. Thirty-two national and six regional VOPEs have formed 25 partnerships to design and implement Peer-to-Peer projects.[4]

The largest evaluation association today is the American Evaluation Association (AEA), with a membership of 7,100, followed by the Canadian Evaluation Society (CES), with about 1,800 members; the Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC) has about 1,600 members; the International Association for Development Evaluation (IDEAS) about 900 members; Australasia 860 members; and the European Evaluation Society (EES) about 550 members.[5] But then the numbers drop substantially. Other national and regional associations generally struggle, with from about 150 members to just a handful (UNDP 2015). Thus, efforts to support VOPEs, which are often easier to create than to sustain, continue to be needed.

The Global Evaluation Agenda, EvalAgenda 2020 (EvalPartners 2016), sets out four key areas where evaluation capacity needs to be strengthened if it is to fully realize its potential in supporting the new development agenda and beyond: an enabling environment for evaluation, institutional capacities for evaluation, the capabilities of individual evaluators, and the links between these three elements. And while all of them are important, it is the third area, concerning the capabilities of individual evaluators, that relates most directly to professionalism.

A strong **enabling environment** is described as one where all sectors of society understand and appreciate the value of evaluation; where evaluation is explicitly recognized or encouraged in national evaluation policies and other governance and regulatory instruments; where sufficient resources are allocated for evaluation at all levels; where evaluation findings are used; and where evaluation receives due recognition as a profession.

Strong **institutional capacities** involve strong VOPEs, as well as government agencies, civil society organizations, academia, and other institutions that generate and share relevant data to support evaluation.

Developing **individual capacities** for evaluation is relevant not only to evaluators but to commissioners and users of evaluation as well. The latter need sound understanding of the value of evaluation, and commitment to using evaluation findings and recommendations. In terms of individual capacities, the goals are to have:

[4] More information on the program and reports from its first projects can be found under P2P on the EvalPartners website, EvalPartners.org.

[5] Membership numbers are from the various association websites, accessed May 2017: AEA, http://eval.org; CES, https://evaluationcanada.ca; the Australasia Evaluation Society, http://aes.asn.au; IDEAS, http://ideas-global.org; and EES, http://europeanevaluation.org. ReLAC membership data are from Rodriguez-Bilella and Lucero (2016).

- Sufficient numbers of qualified evaluators, drawn from a diversity of relevant disciplines, who are available to conduct high-quality evaluations in all countries and all subject areas;
- Evaluators who have the knowledge, skills, and dispositions to make appropriate use of generally accepted evaluation principles, theories, methods, and approaches;
- Evaluators who have integrated the values discussed above, and are culturally sensitive.

But despite the vision presented in this agenda, a lack of clarity and agreement exists today on what it would look like to have national evaluation capacity in the context of the SDGs. UNDP, for example, has contracted for a study to try to further understand what this would entail, and what it would look like if a country had it.[6] The issues are many. For example, does national evaluation capacity mean having a sufficient number of qualified evaluators who are able to conduct high-quality evaluations in all relevant subject areas? Does it mean having the governmental capacity to do cross-cutting evaluations (i.e., evaluations that cross different ministries, so that interrelatedness can be addressed)? Does it mean having government policy makers and parliamentarians who are not only able to use evaluative information, but also consistently do so? Is it about having the capacity to conduct independent evaluations? Inclusive evaluations? Self-evaluations? Rigorous counterfactual evaluations? Complex systems evaluations? Systematic reviews? All the above and more?

THE CHALLENGES OF PROFESSIONALIZATION

It is difficult to build evaluation capacity in a profession that remains fractured, and lacking in agreement on how to define competencies for evaluators. Much has been written about evaluators and development evaluators still lacking professional status and visibility; evaluators not feeling recognized as professionals; and the belief that the lack of control over access to the ranks of evaluators resulting from lack of professional standards (and lack of enforcement of those standards) too often yields poor quality evaluation work.[7]

Is the Problem That Evaluation Is a Young Profession?

Evaluation is often referred to a young discipline. For example, Robert Picciotto, one of the gurus of development evaluation, calls it a "fledgling profession" (Picciotto 2015). Others have called it the "new kid on the block" among the social sciences. Its "youth" is often given as the reason for its struggles with professionalizing, and its general lack of agreement on evaluator

[6] Personal communication with Charles Lufthaus, Universalia, January 13, 2017.

[7] See, e.g., Altschuld and Engle (2015), King and Podems (2014), Morra Imas (2010), and Picciotto (2011).

competencies. But it is interesting to compare two young professions in the United States—school psychologists and evaluators—both of which have developed over similar time frames.

Rossi, Freeman, and Lipsey have credited the boom in evaluation to the demand for knowledge of results that accompanied large public expenditures for major programs in urban development and housing, education, occupational training, and preventive health services following World War II (Rossi, Freeman, and Lipsey 1999). They indicate that major commitments were also made during this time to international programs for family planning, health, and nutrition, and rural development. They conclude that by the end of the 1950s, evaluation research was commonplace.

According to Hogan, we can thank Russia's launch of Sputnik in 1957 and the ensuing space race for the discipline of evaluation (Hogan 2007). The National Defense Education Act poured money into new education projects and programs in math and science, and evaluations were funded to measure the success of the new curricula. The passage of the Elementary and Secondary Education Act of 1965 is commonly considered the birth of modern program evaluation in the United States, because it required evaluation, and thus helped evaluation to emerge as a profession. Once federal monies began to flow, universities began to offer courses in evaluation methods.[8] The Evaluation Research Society emerged in 1976, and so did evaluation journals such as *Educational Evaluation and Policy Analysis, New Directions for Program Evaluation,* and *Evaluation News*. By this time, evaluation clearly had emerged as a distinct specialty field within social science. In 1986, the Evaluation Research Society and the Evaluation Network merged to become the AEA.

Evaluation has been largely practice-based, with the people who do program evaluation coming from many different backgrounds, such as education, sociology, psychology, economics, social work, and public policy. Most evaluation degrees are still awarded out of departments such as education or psychology.

But as described by Stevahn et al. (2005), in the development of competencies for evaluators, the field of program evaluation has been decidedly less than can-do. They indicate that most fields recognized as professions, such as health care, teaching, counseling, and so on, have typically developed competencies for their practices by asking a group of distinguished practitioners—often on behalf of a professional organization—to first generate a category scheme and initial list of competencies, then to institute an expert review process to edit and refine them (Stevahn et al. 2005). The competencies are then made available to professionals in the field so that they can structure training programs for novice practitioners; continuing education programs for experienced professionals; and periodic reviews to update the competencies as theory, research, and practices evolve over time. But this has not happened in the field of program evaluation. Because there has been no

[8] For example, at Western Michigan, the University of Virginia, and the University of Illinois.

standardization, anyone can claim to be an evaluator, and can still do so to this day.

By contrast, the field of school psychologists was recognized as a division of the American Psychological Association (APA) in 1945.[9] But it was a nine-day conference in 1954 with 48 APA participants representing practitioners and trainers of school psychologists that began to advance the profession. Their task was to develop an official position on the roles, functions, and necessary training and credentialing of school psychologists. One of the goals of the conference was to define school psychologists, and the agreed definition was that school psychologists were psychologists who specialize in education, and have specific knowledge of the assessment and learning of all children. Participants at the conference felt that since school psychology is a specialty, individuals in the field should have a completed either a two-year graduate training program or a four-year doctoral program. They also felt that states should be encouraged to establish certification standards to ensure proper training. It was also decided that a practicum experience should be required, to help facilitate experiential knowledge within the field.

The National Association of School Psychologists (NASP) was formed in 1969, and in 1988-89 they moved to a national credentialing system. NASP is linked to state education agencies and to their credentialing boards. The NASP Standards for Training and Field Placement Programs in School Psychology identify the critical training experiences and competencies needed by candidates preparing for careers in school psychology. Ten domains are laid out, with standards in each domain. These standards serve to guide the design of school psychology graduate education. They provide a foundation for the recognition of programs that meet national quality standards through the NASP program approval process. The Standards for the Credentialing of School Psychologists are intended as a model for state education agencies or other state or local entities that employ school psychologists, and have the statutory authority to establish and regulate credentialing for school psychologists' titles and practices. Included are recommended criteria for initial credentialing, consisting of graduate coursework, practicums, and internship requirements, as well as recommendations for credential renewal.

While the world of school psychologists has its own debates, such as whether a doctorate should be required for entry into the profession, there is no debate about professionalism. Access to the profession is clearly highly controlled.

This is not to suggest that the credentialing of school psychologists should serve as a model for global evaluation, where access to training is not equal, and equity is a major concern. Additionally, evaluation is trans-sector: it does not have a history of state licensure, nor is accreditation a goal. But what we can conclude from looking at the development of school psychology as a profession is that the youth of a profession does not necessarily

[9] Much of this section is drawn from the National Association of School Psychologists (NASP) website and from Wikipedia.

correlate with professionalism, or hinder the degree of control over access to the profession through the setting of standards.

Letting a Thousand Flowers Bloom

Today any person or group can create their own set of evaluation competencies. And indeed, that is not only what is happening, but also what is being encouraged. The Global Evaluation Agenda summarizes the general view that "Given widely different cultural contexts and operating requirements, no standard blueprint approach would be appropriate for all VOPEs in all countries" (EvalPartners 2016, 24). Given the widespread contexts and sectors that evaluators cover, it is believed that each VOPE should design its own qualification system within a set of general, internationally accepted, guiding principles. Such principles are currently part of evaluator capabilities framework pilots being implemented by the EES and the United Kingdom Evaluation Society (UKES). These principles address voluntariness, autonomy, legitimacy, pluralism, transparency, equity, and quality assurance.[10]

As recognized by King and Stevahn, there are advantages to letting a thousand evaluation-competency framework "flowers" bloom (King and Stevahn 2015). This provides room for adaptation to unique contexts and content, and it may generate creativity and innovative ideas. The good principles referred to above would not restrict the bloom, and they can help guard against the possibility of elitism and continuing exclusion that is feared, especially in the context of developing countries, if formal qualifications are overemphasized over other indicators of competence, such as on-the-job training or relevant experience (Levin 2015).

But unless there is an agreed-on core of competencies that have some part in competency frameworks, it is hard to see how practitioners will be able to advance the argument that evaluation is indeed a discipline. Without an agreed-on core set of competencies that can be augmented by specialist and context-laden additions, it is difficult to see how the field of evaluation will be able to move toward increased professionalism. While there is value in diversity, more coherence is needed in order to advance the professionalism of evaluation. And without a core set of competencies, there is no sound basis for the exclusion of unqualified practitioners—a basic qualification for any profession.[11]

COMPARING SYSTEMS OF COMPETENCY

Today evaluator competency systems are rapidly being developed and adopted around the world not only in VOPEs, but also in organizations such as the United Nations Evaluation Group (UNEG 2016a, 2016b) and graduate

---

[10] These principles are more fully described in *EvalAgenda2020* (EvalPartners 2016), 84–86.

[11] See Wilcox and King (2014), 3, describing Worthen's nine criteria for judgments of evaluation's professional status.

school programs in universities, for example, Western Michigan University's doctoral program in evaluation. While only two credentialing systems exist—those of the CES (which is peer review–based) and the Japanese Evaluation Society (which is training course and exam–based)—the EES and the UKES are conducting pilots of peer review–based systems. IDEAS, the international VOPE, has been considering whether to also pilot a peer review–based system and/or to investigate other options. Given the more rapid growth of competency programs over credentialing programs, the next questions to consider are how fundamentally different these competency systems really are, and whether any of them have been developed in the context of the SDGs.

Table 5.1 compares the competency domains of five different international, regional, and national associations. The associations have separate ethical standards for their memberships, and some have separate standards for the commissioners of evaluations.

The CES has the only approved and operating peer review–based, credentialed evaluator system.[12] Launched in 2010, it is a voluntary designation, which means that the holder has provided evidence of the education and experience required by CES to be a competent evaluator. Competencies for Canadian evaluation practice (along with ethical standards) are the foundation for the credentialed evaluator program. Their 49 key competencies were placed into five competency domains: reflective practice, which focuses on fundamental norms and values, and awareness of one's evaluation expertise and need for growth; technical practice competencies, which focus on specialized aspects of evaluation; situational practice, which covers the application of evaluative thinking and the contextual circumstances in which evaluation skills are being applied; management practice competencies, which focus on the process of managing evaluations; and interpersonal practice competencies, which focus on "people skills."

To qualify for the designation, applicants must provide evidence of a graduate-level degree or certificate related to evaluation; evidence of two years (full-time equivalent) of evaluation-related work experience within the last 10 years; and indicators of education and/or experience related to 70 percent of the competencies in each of the five domains. As a peer review–based system, applications are reviewed by the CES credentialing board. There are special provisions for those who do not have a graduate degree or certificate, but they do carry additional out-of-pocket costs.

The Aotearoa New Zealand Evaluation Association (ANZEA) established its evaluator competencies in 2011, adding them to a system that included ethical guidelines and evaluation standards for undertaking and

[12] See Kuji-Shikatani (2015). Also, the Japan Evaluation Society has not only developed a competency framework, but has also implemented a certification program that is based on completion of a training program and a passing score on the related exam. As of early 2017, the Eurasian Alliance of National Associations, which includes evaluation associations from the Russian Federation, Ukraine, Kazakhstan, Kyrgyzstan, Tajikistan, and Armenia, was seeking to partner with academic institutions to further professionalization.

TABLE 5.1 **Comparison of competency/capability framework domains of five evaluation associations**

General description	CES	ANZEA	IDEAS	EES	AEA (draft)
Enhances/advances professional practice/ continuous learning	Reflective practice	Reflective practice and professional development	Promoting a culture of learning	Dispositions and attitudes	Professional domain
Knows and applies appropriate design and methods	Technical practice	Systematic evaluative inquiry	Professional foun-dations; monitoring; planning and design; conducting the evaluation	Evaluation knowledge: appreciates role played by evaluation in society; masters antecedents of evaluation quality; understands potential/ limits of evaluation instruments and tools	Methodological domain
Considers/analyzes issues, interests, context, and adapts practice	Situational practice	Contextual analysis and engagement	Professional foun-dations; monitoring; planning and design; conducting the evaluation	Evaluation knowledge: masters antecedents of evaluation quality	Context domain
Conducts/manages evaluations skillfully	Management practice	Evaluation project management and pro-fessional practice	Managing the evaluation	Professional practice: capacity to manage and deliver evaluations	Management domain
Inclusive/commu-nicates effectively and respectfully/ negotiation	Interpersonal practice		Managing the evalua-tion; communicating evaluation findings	Professional practice: displays interpersonal skills	Interpersonal domain

commissioning evaluations (ANZEA 2011). Cutting across four competency domains and approximately 100 competencies are values and cultural competency. These are meant to ensure the inclusion and participation of indigenous groups and all marginalized subgroups. The first domain, contextual analysis and engagement, describes the abilities critical to undertaking analysis of the context; engaging with people as part of developing this understanding; and identifying the people, skills, knowledge, and experience needed to carry out the evaluation. The second domain, systematic evaluative inquiry, describes the knowledge, skills, and abilities required to undertake a systematic evaluative inquiry. The third domain, evaluative project management and professional evaluation practice covers the competencies needed to manage an evaluation in a professional manner. The fourth, reflective practice and professional development, includes competencies that support the development of the evaluation practitioner and the profession.

In 2012, IDEAS also approved a set of professional competencies. IDEAS is the only association for individual development evaluators, and it can proudly state that its competencies were developed by an international volunteer group from 40 countries, and ratified by the IDEAS membership, which spans at least 105 countries (Morra Imas 2010). Competencies were developed for those who conduct evaluations as well as for those who manage evaluations but do not conduct them directly. For those evaluators conducting evaluations, 25 competencies across six domains were identified. Additional supporting documentation breaks the competencies down even further. This is the only one of the five competency frameworks reviewed that identifies monitoring skills and capabilities as a domain. Additionally, unlike with the other competency systems, technical skills are broken into stages. Adapting to and knowing the context and the culture—what other systems refer to as situational practice or context—is incorporated by IDEAS into professional foundations; planning and design; and conducting the evaluation. Interpersonal practice also overlaps several domains. IDEAS treats all the competencies as core competencies.

The EES has attempted to make a distinction between a capabilities, or input-based, system such as their own, and ANZEA's and other outcome-based competency systems, such as the CES system. Outcome-based systems require evaluators to demonstrate their competencies: they are in effect testable, or results-based. Input-based systems are viewed as having a more deliberate learning orientation that focuses more on capabilities than on the demonstration of skills. While this distinction may not be quite clear at this point, and while the framework is being used to implement the Voluntary Evaluator Peer Review Pilot, it is certain that this is the only framework that stresses evaluator dispositions and attitudes, rather than reflective practice or professional development. Other models tend to make evaluator dispositions and attitudes a focus of separate ethical standards, or they include a competency on compliance with ethical standards.

After years of discussing competencies (Stevahn et al. 2005), in 2015, the AEA began to formally develop a set of competencies for its membership (Altschuld and Engle 2015). Several drafts have been produced and reviewed by the membership since that time, with review continuing into

September 2017. The specific goal of the AEA is to develop the general competencies that every evaluator or team of evaluators should have, regardless of context. As of September 2017, the competencies remained in draft form, with 44 competencies comprising the five domains. AEA continues to debate certification and credentialing.

SURPRISING COMMONALITIES

Looking across the five frameworks, at least in terms of the domains, there is more consistency than might be expected. While the specific words used may differ, five domains seem to be central to all of these systems: reflective practice; professional or methodological skills; contextual understanding; evaluation management; and interpersonal communication. A few domains are unique, such as the IDEAS identification of monitoring practice, or EES's dispositions and attitudes.

More variation seems to exist in the specification of competencies, and in their placement in domains. Some associations specify more methodological evaluator competencies, while others focus on competencies in the interpersonal domain. Still, this comparison suggests that identifying and agreeing on five or six core domains, and core competencies within each domain that skilled evaluators in various contexts should have, is a realizable goal. The core piece could then be added to and adapted for evaluators working in different specific contexts, but the core would remain the same.

None of these five systems have been developed with the SDGs or MDGs as a driving force. Most acknowledge that they will require review and revision from time to time in order to remain current. If the SDGs, and the commitment to them, are to be taken seriously, then they should provide the imperative for a review against the core competencies required by those who seek to evaluate the SDGs. It may be only a dream now, but the SDGs could provide the impetus to move forward toward the professionalization of evaluation in a more directed way.

CONSIDERING CORE COMPETENCIES IN THE CONTEXT OF THE SDGS

Thinking through core competencies in the context of the SDGs will likely be a three-step process. What is needed first is a review and agreement on the handful of core domains and concomitant competencies that are most important for skilled evaluators to have, whether they are specialists in empowerment evaluation or randomized designs, HIV/AIDS evaluation experts, or country evaluation specialists. These are the base competencies that those who call themselves evaluators should have, even though they may also specialize in particular evaluation methods, sectors, or countries. Like any set of competencies, these would not be set in stone, but would have provisions for periodic review and renewal as the field evolves. This step in itself would advance the professionalism of the field.

Second, as discussed earlier, is the envisioning of what it would look like at the national level, for countries to have the capacity to evaluate the

SDGs. The third would be to add to the core competencies and extend them to the SDG context as needed.

This will not be relevant for all evaluators. Not all evaluators work on an everyday basis in the direct context of the SDGs. But it would be a critically important paradigm for the many evaluators who are struggling with the SDGs. For example, working across sector boundaries might become a core SDG competency under the interpersonal domain. Mastery of complexity theory and systems approaches might be deemed important SDG competencies under professional or methodological practice; or new uses of technology for better data capture; or the use of big data. Some key questions will always need to be asked—for example, is this a necessary competency? or is it teachable?, as suggested by King and Stevahn (2015).

What are some of the advantages of professionalizing evaluation in the context of the SDGs? By creating agreement on what the evaluation of the SDGs entails, and delineating the core competencies necessary to undertake it, clear core standards would enable evaluators to work across geographical boundaries. This could level the playing field in the sense of knowing what the expectations are, and enabling the targeting of evaluation training programs where they are lacking, and where they are most urgently needed. Flexibility could be retained for noncore competencies, so that customization is still possible. Two of the largest benefits might be increased comparability of evaluation findings from shared methods and approaches, and an increased quality of evaluations.

How to start? That the convening power exists today to accomplish this goal is clear. EvalPartners, in partnership with IOCE, IDEAS, UNEG, and others, could continue a series of global multistakeholder consultations, whether face-to-face, virtual, or some combination of the two. This has already started to some extent, with the Third Global Evaluation Forum, held in April 2017 in Bishkek, Kyrgyz Republic, and organized by IOCE, EvalPartners, UNEG, the Global Parliamentarian Forum for Evaluation, the Kyrgyz Monitoring and Evaluation Network, and the Eurasian Evaluation Network, with the support of the Kyrgyz government. This forum brought together some 150 delegates, representing governments, parliaments, development partners, foundations, the private sector, universities, civil society, and the evaluation community, to advance implementation of the SDGs through review and implementation of Eval2020. Much could be gained by involving larger groups of evaluators, taking advantage of everyone's need to understand and advance evaluation of the SDGs; and with it, the opportunity for professionalization of the field. As SDG competencies are developed and agreed upon, ready access to high-quality professional training opportunities on SDG evaluation can follow.

REFERENCES

Altschuld, J.W., and M.E. Engle, eds. 2015. *Accreditation, Certification, and Credentialing: Relevant Concerns for U.S. Evaluators.* New Directions for Evaluation, No. 145. San Francisco: Jossey-Bass.

ANZEA (Aotearoa New Zealand Evaluation Association). 2011. *Evaluator Competencies*.

Berliner, Tom. 2015. "Can the SDGs be Achieved by 2030?" *Independent Research Forum* blog.

Catsambas, T.T. 2016. "The 5 Practices of SDG-Ready Evaluation." *Evaluation* blog.

EvalPartners. 2016. *EvalAgenda 2020. Global Evaluation Agenda: 2016–2020*.

Hogan, R.L. 2007. "The Historical Development of Program Evaluation: Exploring the Past and Present." *Online Journal of Workforce Education and Development* 11 (4): 1–14.

King, J.A., and D. Podems. 2014. "Introduction to Professionalizing Evaluation: A Global Perspective on Evaluator Competencies." *Canadian Journal of Program Evaluation* 28 (3): vii–xv.

King, J.A., and L. Stevahn. 2015. "Competencies for Program Evaluators in Light of Adaptive Action: What? So What? Now What?" In *Accreditation, Certification, and Credentialing: Relevant Concerns for U.S. Evaluators*, J.W. Altschuld and M.E. Engle, eds., 21–37. San Francisco: Jossey-Bass.

Kuji-Shikatani, K. 2015. "Credentialed Evaluator Designation Program, the Canadian Experience." In *Accreditation, Certification, and Credentialing: Relevant Concerns for U.S. Evaluators*, J.W. Altschuld and M.E. Engle, eds., 71–85. San Francisco: Jossey-Bass.

Levin, R.M. 2015. "Professionalizing Monitoring and Evaluation for Improved Performance and Integrity: Opportunities and Unintended Consequences." Presentation at the SAMEA Conference.

Morra Imas, L.G. 2010. "The Movement for Global Competencies for Development Evaluators." Paper presented at meeting of the Canadian Evaluation Society, Victoria, B.C.

Picciotto, Robert. 2011. "The Logic of Evaluation Professionalism." *Evaluation 17* (2): 165–80.

———. 2015. "What Is Evaluation?" Lunch presentation, International Program for Development Evaluation Training.

Rodríguez-Bilella, P., and M.A. Lucero. 2016. "Evaluation as a Global Phenomenon: The Development of Transnational Networks." In *The Future of Evaluation: Global Trends, New Challenges, Shared Perspectives*, R. Stockman and W. Meyer, eds., 66–80. London: Palgrave MacMillan.

Rossi, P.H., H.E. Freeman, and M.W. Lipsey. 1999. *Evaluation: A Systematic Approach*. 6th edition. Thousand Oaks: Sage Publications.

Schwandt, Thomas, Zenda Ofir, Dorothy Lucks, Kassem El-Saddick, and Stefano D'Errico. 2016a. "Evaluation: A Crucial Ingredient for SDG Success." IIED Briefing, International Institute for Environment and Development, London.

Schwandt, Thomas, Zenda Ofir, Stefano D'Errico, Kassem El-Saddik, and Dorothy Lucks. 2016b. "Realizing the SDGs by Reflecting on the Way(s) We Reason, Plan and Act: The Importance of Evaluative Thinking." IIED Briefing, International Institute for Environment and Development, London.

Stevahn, L., J.A. King, G. Ghere, and J. Minnema. 2005. "Establishing Essential Competencies for Program Evaluators." *American Journal of Evaluation* 26 (1): 43–59.

UN (United Nations). 2015. "Transforming Our World: The 2030 Agenda for Sustainable Development." A/RES/70/1. UN General Assembly, New York.

UNDP (United Nations Development Programme). 2015. *Towards a Baseline Study: Insights on National Evaluation Capacities in 43 Countries*. New York: UNDP.

UNEG (United Nations Evaluation Group). 2016a. *Evaluation Competency Framework*. NY: UNEG.

———. 2016b. "Professionalization of Evaluation." Concept Paper. UNEG, NY.

———. 2016c. "Evaluation in the SDG Era: Lessons, Challenges and Opportunities for UNEG." Working Paper. UNEG, NY.

White, Howard. 2015. "Toward Evidence-Informed Policies for Achieving the Sustainable Development Goals." *Evidence Matters* 3ie blog.

Wilcox, Y., and J.A. King. 2014. "A Professional Grounding and History of the Development and Formal Use of Evaluator Competencies." *Canadian Journal of Program Evaluation* 28 (3): 1–28.

Chapter 6

From Evaluation Capacity Building to Evaluation Capacity Development – A Paradigm Shift

Michele Tarsilla

Abstract. *In acknowledging the current limitations of contemporary evaluation capacity-building practice, and in an effort to promote an innovative and equity-focused contribution to the current discourse on evaluation capacity, this chapter suggests a new framework for conceptualizing, planning, implementing, and assessing the results of capacity development (as opposed to capacity building) in evaluation more effectively in the future. The first part of the chapter proposes a new definition of capacity in evaluation, and encourages readers to embrace and adopt the more encompassing term of "evaluation capacity development" as opposed to that "evaluation capacity building." The second part offers funders and planners an overview of those contextual and process-related factors that need to be taken into account in order to enhance the effectiveness of capacity development activities and programs. The third part highlights the specific and innovative contribution of the International Development Evaluation Association (IDEAS) to the ongoing discourse on evaluation capacity.*

Michele Tarsilla, International Development Evaluation Association, mitarsi@hotmail.com.

Often associated with the delivery of short-term training and technical assistance funded by international agencies, evaluation capacity building (ECB) has fallen short of its intended objectives in many different development and humanitarian contexts (Tarsilla 2014a). Despite the copious resources allocated by international development organizations to enhance the capacity of low- and middle-income countries to evaluate the performance and impact of their development programs and policies, most ECB activities on the ground have failed on multiple fronts. They have not been able to target a sufficiently large critical mass of individuals and organizations that could truly foster systemic change in the uptake and use of evaluation. The agencies that fund evaluation capacity-building programs in the Global South have not coordinated with each other as closely and systematically as they should have and, in so doing, have undermined their own ability to foster a more efficient and strategic use of resources. Finally, time and energy have been focused on the measurement of short-term effects, while the quest for long-term results has largely remained elusive.

Well aware of such weaknesses, and in response to the need expressed by many actors for the roll-out of more innovative ECB strategies, the International Development Evaluation Association (IDEAS) is promoting a paradigm shift in contemporary thinking and practice in the area of evaluation capacity development (ECD) (Tarsilla 2012). In particular, it is calling for a shift from the current focus on short-term training activities to the adoption of ECD strategies that are more contextually relevant and are better geared toward equitable, systemic and sustainable learning in evaluation.

EVALUATION CAPACITY BUILDING: KEY ISSUES AND LIMITATIONS

It is well understood that having individual practitioners and organizations' staff participate in two- or three-day workshops on evaluation theories and methods cannot foster dramatic change either in the quality of evaluation practice or the use of evaluation products. However, despite this, most ECB interventions, supported by both national and international funders in many different countries over the last two decades, have consisted mainly in the implementation of a finite number of activities. Such reification of capacity building, illustrated by the tendency to equate capacity building with training, as well as the tendency to implement evaluation workshops in a vacuum—that is, without accurate knowledge of how power and resources are distributed and contested at the local level—is indeed one of ECB's main weaknesses. This phenomenon, which has had egregious effects on the way international and national funders have planned and budgeted for in this area of development in the past, has been so prevalent that the meaning of ECB has been watered down, and its potential significantly compromised.

A second limitation in the way ECB has been conducted in the past is the dismissal of organizational processes in the planning, implementation, and evaluation of capacity-building efforts. While too much focus has been given to advancing the technical skills of individuals and organization staff, ECB planners and workshop facilitators have often failed to assess and act

upon the environment in which their target groups are operating. In particular, ECB planners and implementers have not systematically addressed the following as part of their capacity needs assessment:

- Environmental factors that influence the use of evaluation (e.g., the type of political system in place at the national and subnational levels; existing governance and accountability norms; and the degree of openness to accept failure/s and learn from them)
- Institutional or organizational processes that either benefit or hinder the evaluation function (e.g., lines of reporting and communication across different levels of a governmental agency, or data quality assurance regulations within the national statistics office)
- The quantity and quality of incentives available to conduct and use evaluation (e.g., the systematic publication of evaluation reports on a public portal, and/or the practice of organizations' executives to develop a management response in reaction to the recommendations included in an evaluation report)

Given inadequate understanding of ECD ecology, most funders and planners have failed to get many of the relevant actors from either the public or private sectors involved, either as partners or beneficiaries, in their past ECD efforts.[1] Evaluation training programs, for instance, are generally aimed at a limited number of evaluation technical officers from one or more organizations without the strategic involvement of their supervisors and directors. Furthermore, numerous ECB interventions provide participants with evaluation toolkits and checklists but they often dismiss the environmental factors that influence the adoption and use of such knowledge product—what I have defined as the "political economy" of ECD.

A third limitation has been the more or less inadvertent perpetuation of the old development paradigm, according to which donors' needs and interests prevail over anybody else's. For more than a decade, ECB activities have been geared toward increasing the level of knowledge, skills, and attitudes of project staff in the field with the primary de facto objective to enhance regular results reporting to funders (e.g., on a quarterly basis). What has been particularly fallacious is the assumption that retrofitting existing practices within established boundaries identified by donors would enable empowerment and social change. Unfortunately, this donor-centric strategy, which I tend to classify as "functional evaluation capacity building," or F-ECB, not ECD, has gradually become the norm (Tarsilla 2014a). In one case, four small cultural organizations that I worked with in the Democratic Republic of Congo had strategic objectives that, as spelled out in the plans and logical frameworks formulated for them by an international funder, were out of sync not only with their organizations' own vision, but also with their country's

[1] Public sector actors would include, among others, staff in ministries and members of parliaments. Private sector actors would include, among others, training institutes and consulting firms.

national cultural policies. It was very disheartening to discover that—according to the contract signed with the funder—I needed to enhance the capacity of these four organizations to measure the attainment of their objectives, which had been imposed from the outside, by using a number of indicators that the organizations in question did not really understand, or assign any credit to. In particular, it was very difficult to talk to them about logic models, theory of change, and rigorous evaluation designs, especially given the fact that the impact indicators showing at the top of the logical framework template—which had been distributed by the funder—rested on the assumption that grantees would be in a position to effectively measure the extent to which some of their activities, which were targeting less than 20 participants per year, had contributed to improved attainment of two of the Millennium Development Goals in the whole country. Unfortunately, this example shows once again that ECB programs often provide participants with the knowledge and tools that facilitate timely reporting to funding agencies, but rarely foster true organizational learning and increased results-oriented agency.

HOW TO OVERCOME EXISTING LIMITATIONS WITH THE NEW DEFINITION OF CAPACITY IN EVALUATION: FROM EVALUATION CAPACITY BUILDING (ECB) TO EVALUATION CAPACITY DEVELOPMENT(ECD)

ECB in International Development: Key Assumptions and Real-World Considerations

The evaluation policies and strategic evaluation plans currently in use among several development agencies around the world are predicated on the assumption that international development evaluation serves two primary functions (GIZ 2013; Norad 2006; Sida 2007; UNESCO 2015; USAID 2011). The first is to enhance the accountability of those who manage and implement international development projects, especially vis-à-vis their respective funders and expected beneficiaries (Wiesner 1997). The second is to foster learning among those who commission, manage, conduct, and use evaluation, on what works well and what needs to be improved in international development projects and programs (Argyris and Schön 1996; Bamberger 2009; Pasteur 2006; Rist, Boily, and Martin 2011; Solomon and Chowdhury 2002). Based on such assumptions, any activity aimed at strengthening the evaluation function—locally, nationally, or globally—should, therefore, be able to contribute to strengthening both the performance and the effectiveness of international development projects in a variety of countries. However, this is easier said than done.

Typically, funders and international development agencies attempt to strengthen the evaluation function by developing the technical skills of local development practitioners. However, such strategies do not always translate into stronger development effectiveness. One reason for this is the lack of a genuine evaluative culture—that is, the systematic conduct of evaluation, and the use of findings for decision making—which often results from the limited ability of ECB to foster ownership and inclusiveness of evaluation processes.

The scenario, however, is not as bleak as it would first appear. There are several examples of countries that have put monitoring and evaluation (M&E) systems into place that are both prospering and serving accountability and learning purposes. Chile, Colombia, Malaysia, Mexico, South Africa, and Sri Lanka provide good illustrations of how the creation of a supportive environment to enhance the evaluation functions at both the organizational level (Stevenson et al. 2002; Wijayatilake 2011) and the institutional level can foster the development of a strong national evaluative culture (Boyle, Lemaire, and Rist 1999; Mayne 2008; Trochim 2006). One feature that all of these countries have in common is the buy-in of different stakeholder groups both within and outside the national government into the discourse on evaluation, and their subsequent involvement in all of the related processes. Such success stories, though, have not been capitalized on in the international development arena as frequently as they should, or could, have been. The result is that the pursuit of non-inclusive targeting strategies has hindered the success of ECD interventions in many countries around the world. The simultaneous involvement of actors operating both within and outside of national governments as part of an ECB intervention has been very rare: this is also a result of the rigidity featured by the mission and the scope of work among the majority of funding agencies. On the one hand, for over a decade bilateral donors and philanthropic foundations have been able to fund a plethora of initiatives and programs specifically aimed at strengthening the knowledge and skills of stakeholders supplying evaluation services (e.g., representatives from academia, the private sector, and nongovernmental organizations) (OECD 2006). On the other hand, multilateral agencies, such as the United Nations and the World Bank, have allocated the largest share of ECB resources to national governments; that is, the agents that for the most part demand evaluation services.

ECB or ECD: Does the Terminology Really Matter?

Far from becoming rhetoric in a purely academic exercise, an exploration of the language used to describe ECB would be particularly useful to enhance future programming in this area. In particular, reviewing the central attributes, modalities of implementation, and evaluative variables of ECB would be beneficial. The same is true for another popular term used in the international development arena: evaluation capacity development. An in-depth review of ECD appears even more relevant than that of ECB, given that capacity development—in evaluation as well as in a variety of other fields—is not the only "missing link in development" (World Bank 2005, 24). Even more importantly, it is part of the overall goal of development cooperation (Fukuyama 2004).

A certain confusion or lack of consensus exists over the meaning of both terms, and has been attested to by a stream of peer-reviewed literature produced by evaluation scholars. Among the most recent contributions on this topic, Bohni and Attström's (2011) appears particularly relevant. According to these Danish authors, more serious reflection and debate on the distinction between ECD and ECB is needed, as it would allow the addressing of four main issues affecting the practice of evaluation in a number of

countries: the widespread conceptual pluralism in the area of ECB and ECD; the increased number of discordant opinions regarding the purposes of ECB; the lack of a comprehensive empirical base for most ECB and ECD models; and the relatively greater focus on the approaches implemented in tackling ECB rather than ECD.

When referring to capacity building, Morgan, one of the most prolific authors on capacity, has defined it as:

> ...a risky, murky, messy business, with unpredictable and unquantifiable outcomes, uncertain methodologies, contested objectives, many unintended consequences, little credit to its champions and long time lags. (Morgan 1998, 6).

Likewise, in defining capacity development, Lusthaus, one of the most well-respected Canadian experts in institutional evaluation and change, has described it as follows:

> ...a concept still in its infancy. Its definition is still forming. Research describing how people use the concept is sparse. So is research, which tests its assumptions and predicts its consequences. There are few evaluations of projects that are claiming to use approaches to capacity development. (Lusthaus et al. 2002, 34)

The discourse on capacity development (including over its definitions) has continued over the years. However, it has traditionally been dominated by the voices of northern scholars. Only recently have researchers and academic institutions in the Global South become engaged in this area of work. Among some of the most recognized actors who have contributed to advancing capacity development-related definitions are the following:

- The African Union and the New Partnership for Africa's Development, which in 2010 introduced an African Capacity Development Strategic Framework (AU and NEPAD 2012)
- The African Capacity Building Foundation (ACBF 2016)
- The Capacity Development for Development Effectiveness (CDDE) forum in Asia[2]
- The Task Team on South-South Co-operation (TT-SSC) in Latin America[3]

The Main Attributes of ECD

Rather than being a purely semantic issue, the distinction between ECB and ECD appears all the more relevant due to the unique political and ideological connotations ascribed to each of the two terms. As confirmed by the

[2] http://www.southsouthcases.info/casosasia/caso_14.php.

[3] https://www.oecd.org/dac/effectiveness/46080462.pdf.

increasing use of the term "evaluation capacity development" in some of the evaluation offices of international organizations, such as the World Bank and GIZ, ECD is characterized more *consistently* and *intentionally* by a stronger emphasis on inclusiveness, flexibility, development result focus, and context responsiveness to an already existing evaluation capacity. An Institute of International Development Studies report commissioned by UNESCO draws an important distinction between capacity building and capacity development that appears particularly useful for the discussion in this chapter (Ortiz and Taylor 2009). When applied to the evaluation context, the report seems to corroborate and amplify the differences between ECB and ECD. As discussed in one of the report's most salient passages, capacity building and capacity development are not described as simply different terms, but rather as two opposite development paradigms:

> Much of the capacity development literature stresses the fact that development is already happening before the arrival of any project, donor, program or initiative, and not to recognize this as an irresponsible error and ultimately a precursor to an ineffective use of resources. Too many donors and executing agencies are determined that their projects be executed in any event, yet when those projects are severely out of tune with the development processes already in motion, they are likely to fail. They fail because:
>
> a) Capacity development programming that does not recognize development in motion is quite literally a foreign object; that is, it pushes ideas that aren't likely to take hold because they are out of step with local realities;
>
> b) They do not build on momentum; that is, positive development initiatives and processes already in motion;
>
> c) The motivation needed to take forward a strategy that does not fit will in turn require a push strategy to convince people to carry it out. Even when the appropriate incentives are in place, true motivation will be dubious because participation will likely be led by the possibility of short-term gain. The fundamentals required for sustainability will be lacking and therefore the project activities and desired behavior changes are unlikely to develop deep roots" (Ortiz and Taylor 2009, 26)

Based on such foundational work, and following global research on the understanding of ECB and ECD among evaluation planners, managers, and practitioners in many different countries, the need for a new definition of ECB and ECD becomes apparent (Tarsilla 2012). However, rather than seeing the two terms as opposite, the two definitions that I came to develop after my exchanges with more than 150 practitioners in over 40 countries situate themselves along an ECB-ECD continuum, where ECB generally accounts for one of the preliminary phases of a broader and long-term ECD strategy.

Evaluation Capacity Building: A New Definition

The new definition of ECB, which was developed toward the end of a long series of consultations with practitioners around the world, reads as follows:

> A necessary (but not sufficient) condition for ECD to take place. ECB mainly consists of a vast array of trainings and coaching activities (some of which are short-term in nature) aimed at building capacity, especially where capacity is either very low or thought not be in place yet, among a discrete number of individuals working either for or within organizations and/or institutions that develop, commission, manage, conduct and/or use evaluation. Although it is an integral component of most national and international development projects today, ECB has often been viewed (especially outside of the United States) as a relatively limited accountability-driven tactic rather than a full-fledged strategy aimed at attaining organizational learning as well as other developmental objectives. As a result, ECB scope and modalities of delivery have often been considered too narrow. (Tarsilla 2012)

Evaluation Capacity Development: A New Definition

ECD, on the other hand, is defined as:

> A process consisting in both the integrated enhancement and maintenance over time of: (a) Individuals' knowledge, skills and attitudes; (b) Organizations' capabilities; and (c) Institutions' readiness; toward contextually relevant planning, management, implementation, and use of evaluation at any level-global, regional, national or sub-national. More specifically, ECD is aimed at both individual and collective transformational learning in the pursuit of three primary goals: strengthening the technical quality and ownership of national evaluation processes; enhancing the local authenticity and cultural appropriateness of evaluation approaches, methods and tools used in-country; and increasing the use of evaluation findings as a way to improve development interventions in a variety of sectors. (Tarsilla 2012)

In order for ECD to be successful, it is critical that ECD strategies be implemented either in a simultaneous, or an intentionally sequenced fashion. ECD-savvy strategies (such as the ones adopted by IDEAS) are specifically aimed at promoting the conditions that support ECD among a variety of actors operating in two different spheres (both within and outside of national government), and characterized by different functions (operational, and policy or decision making) and roles (both consumers and providers of evaluation). ECD strategies consist of a combination of short, medium, and long-term activities (including training, mentoring, coaching, peer exchange, and the creation of evaluation units). Otherwise, ECD appears to be a systemic and adaptive process rather than the combination of stand-alone activities aimed at enhancing capacity at the individual, organizational, or institutional levels. As stressed by the new definition provided above, ECD emerges as an endogenous process that builds upon existing levels of knowledge, skills, and attitudes (individual), capabilities (organizational), and readiness (institutional) either simultaneously or sequentially, and in a variety of contexts (global, regional, national, and subnational), as opposed to building from scratch.

ENVIRONMENTAL AND PROCESS-RELATED FACTORS INFLUENCING THE OUTCOMES OF CAPACITY DEVELOPMENT (ECD) ACTIVITIES AND PROGRAMS

ECD: Gaps in Targeting and Possible Solutions

Activities aimed at strengthening technical capacity within national ministries and central planning agencies (Compton, Baizerman, and Stockdill 2002) have enhanced the knowledge of evaluation within the governmental sphere, but have not necessarily contributed to the development of skills needed to either formulate key evaluation questions or use evaluation findings (Bamberger 2009; OECD 2006). Furthermore, activities aimed at enhancing national evaluation capacity have rarely been customized to the specific functions (operational or strategic) and roles (commissioners, implementers, policy makers) of individual officers operating within the government, and have instead favored the implementation of the same standardized approach at several levels within the government, as if it were a monolithic block. Addressing the limitations of current ECB targeting is all the more relevant, as the currently biased allocation of funding between governmental and non-governmental actors has three primary consequences.

First, it has hindered the mainstreaming of evaluation at a more systemic level, as predicated by a number of studies, including a recent work funded by the U.K. Department for International Development (Gaventa and Barrett 2010). Change happens through multiple types of citizen engagement: not only through formal governance processes, even participatory ones, but also through associations and social movements that are not created by the state. Strengthening these broader processes of social change and their interactions can, in turn, create opportunities for state reformers to respond to demands, build external alliances, and contribute to state responsiveness (Gaventa and Barrett 2010).

Second, the identification of individual evaluation champions within host governments that are characterized by high employee turnover has not always contributed to either the uptake of an evaluation culture or the sustainable promotion of the use of evaluation findings in other sectors (Lennie 2005). There are certainly some good examples of the contributions of national evaluation champions. However, the tendency among politicians to cater to their constituencies' needs and interests regardless of what the available evidence suggests confronts ECD planners and implementers with a real hurdle to overcome.

Third, the greater focus on the demand for evaluation, which some authors critically refer to as "elite domination" (Fung 2003, 340), has ignored the potential contribution of evaluation "suppliers" (e.g., national evaluators), and has not sufficiently leveraged their wealth of knowledge and practical experience during the undertaking of evaluations. That notwithstanding, ECD targeting is already gradually evolving, as attested to by the support provided by such initiatives as EvalPartners and the strengthening of voluntary organizations of professionals in evaluation (VOPEs) over the last five years.

ECD Processes: What's New Compared to Past ECB Practices?

As per the new definition of ECD, the focus of ECD is not on either activities or products (e.g., a training on mixed methods, or the timely submission of a midterm review to the project funder), but rather on processes, interactions, incentives, leadership, organizational learning, and organizational development. Furthermore, ECD is characterized as a particularly inclusive process that is able to respond to the continually emerging needs and interests not simply of individuals, organizations, or institutions; but rather of individuals situated within organizations, and institutions interacting with each other, both in the governmental and nongovernmental spheres. Likewise, rather than resting upon linear and mechanistic planning, ECD is understood and defined as a process grounded in both a realistic understanding of the world's complexity, and the need to adopt more flexible and iterative planning processes. However, my analysis of contemporary ECB practice shows that the latter has focused on developing individual technical skills—how to write evaluation terms of reference, or how to develop sampling strategies—rather than on developing organization-level capabilities and institutional readiness, based on relatively linear and results-based planning processes (Tarsilla 2012). Moreover, in contrast to the descriptions of capacity development as an incessant endogenous process in the peer-reviewed literature, the way the term ECB has been understood suggests that it rests on the main assumption that in-country capacity is static (you either have it or you don't), and that targeted interventions, often funded from external development partners, are the most effective. As a result, ECB does seem to fail to recognize the inherent institutional processes and social dynamics of the settings where its activities are being implemented.

Other Contextual Factors Influencing the Outcomes of ECD Activities and Programs

If one takes the organization as the main unit of analysis of any reflection on ECD,[4] then it is relevant to consider those unique distinct organizational features (its structures and processes) that are likely to affect the outcomes of an ECD program. The adequate consideration of organizational infrastructures and underlying dynamics is, therefore, critical to the success of any ECD strategy. For each of the relevant factors identified in the left column of table 6.1, a series of ECD strategies are recommended in the right column.

Based on a review of the organizational factors listed in the table, a key conclusion is that, despite the size of one's own organization, the planning of any ECD program cannot overlook the context in which that organization operates. As harmonization, relevance, and ownership are some of the principles that any sound ECD program should feature, it could be useful to link organizations targeted by ECB efforts with each other as if they were

[4] Most individual evaluation practitioners work either within or for one or more organizations.

partners and not simply grantees sharing the same donor, so as to allow an alignment of their internal M&E systems.

It is important to note that what is being advocated here is not the development of a cookie-cutter approach, whereby individual organizations share exactly the same set of indicators or evaluation strategies: after all, creativity and flexibility to adjust to emerging changes are two typical features of successful ECD programs. By promoting ECD alignment, what is being suggested instead is to encourage organizations to find common and cost-effective solutions to their information needs and operational questions. During the planning of an ECD program, for instance, each organization with a vested interest in ECD could look at what tools similar organizations have used to measure a certain construct that they are interested in measuring and—for the sake of avoiding replication and the wasting of resources—might adopt those very same existing tools. Besides freeing up resources for the development of other tools, or the set-up of an information management system within one's own organization, for example, ECD alignment would foster the use of identical indicators and, as a result of increased data aggregation, the availability of data that could then inform ECD-related decision and policy making. This is even more relevant if such alignment not only takes place locally, but also at the provincial, district, regional, and national levels.

Likewise, it is important that ECD strategies acknowledge more effectively what the real interests and needs of any organization are, independently of the donors' interests and needs. With that in mind, the following key recommendations should be taken into account during the development of a new ECD strategy:

- Although it is tempting to introduce radical changes within the realm of organizational practices when ECD programs are being implemented, it is critical that donors and contractors implementing ECD strategies on the ground recognize the speed of local organizations to "digest" new evaluation methods and tools.
- Development organizations should understand that assigning a prominent role to funders' evaluation requirements and needs, and building upon them to design an ECD program, is a conventional form of evaluation capacity building. For evaluation capacity development to occur, the centrality of the organizations (each with their own interests and needs) within the system where they are operating needs to be recognized.
- A broader and more systemic targeting of ECD is needed. Two new possible scenarios could be envisaged. On the one hand, funders and implementing organizations should promote the conduct of evaluation awareness-raising among actors who, despite not being directly targeted by the ECD intervention, still gravitate within the system where the latter is being implemented. On the other hand, funders and implementing partners should ensure more active involvement of the entities working in both the private and public sector anytime a large-scale ECD program sponsored by a national

TABLE 6.1 **How an organizational diagnostic could contribute to effective ECD programming**

Key organizational identity traits	Considerations to integrate into an ECD strategy
Historical development (informal development, formalization, expansion, regional consolidation, transition to national ownership, stagnation/implosion, self-reflection, nominal/effective revitalization)	▪ Learn about the history of the organization(s) targeted by your ECD strategy (e.g., key milestones, original founders, individuals promoting and/or challenging transformations within the organizations in question) ▪ Make sure to conduct a stakeholders' mapping, as well as a political economy analysis to assess how power and resources are distributed in the context of the organization(s) where the capacity is expected to develop further as a result of your intervention
Organizational development phase (pioneer, differentiated, integrated) (Ubels 2010)	▪ Assess the extent to which the staff of the organization(s) targeted by your ECD strategy mainly rely on one only leader to find their direction; or if they are driven by clearly articulated organizational policies and job descriptions ▪ In the latter case, explore to what extent evaluation tasks and responsibilities could be integrated into the existing processes and procedures
Membership diversity (low, medium, high)	▪ Learn about the staff making up the organization(s) targeted by your ECD strategy ▪ Assign staff members to different groups according to their specific role and responsibilities (e.g., top leaders and decision makers; managers; technical officers) ▪ Be sure to combine activities aimed at the whole staff with more specific strategies tailored to the needs and interests of each one of the identified subgroups ▪ Try to learn about past training programs offered to each of the identified subgroups so as to build on the examples and language used in past training events
Compliance with internal government rules (low, medium, high)	▪ Look for any organizational and performance audits that have been conducted in relation to the organization(s) targeted by the ECD strategy ▪ Identify organizational deficiencies observed in the past with respect to conformity with the established organizational procedures

(continued)

TABLE 6.1 *(continued)*

Key organizational identity traits	Considerations to integrate into an ECD strategy
Degree of internal networking (low, medium, high)	▪ Reconstruct the lines of communication and reporting among the staff of the organization(s) targeted by your ECD strategy ▪ Liaise with the management information system officer (if available) in the organization(s) targeted by your ECD strategy, to better understand the flow of information, both bottom-up and top-bottom
Resilience (low, medium, high)	▪ Understand how the organization(s) targeted by your ECD strategy has been able to respond to external challenges (even those threatening the survival of the organizations in question) in the past ▪ Highlight the self-defense mechanisms, values, practices that have proved instrumental in allowing the organization(s) to stay abreast of difficulties encountered ▪ Measure to what extent the organization(s) has/have been able to absorb, adapt, and transform
Leadership type (concentrated, decentralized, shared) (Ubels 2010)	▪ Meet with the leader(s) of the organization(s) targeted by your ECD strategy, and try to assess the degree to which their decision-making processes are participatory and inclusive of all staff perspectives ▪ Organize structured conversations with such leaders before implementation of the strategy begins, and try to learn what their respective frameworks of reference are (this might include assessing the type of literature, or the sources from which they draw the information that is informing their decisions)
Ownership (low, medium, high)	▪ Explore the extent to which the organization(s) targeted by your ECD strategy has/have actively participated in the design of the policy or project in relation to which your ECD strategy is being undertaken ▪ Measure the degree to which the different subgroups identified with the organization(s) in question have contributed to, and are still contributing to, the development, implementation, and evaluation of your ECD strategy ▪ Identify the opportunities for scaling up the strategy (this will include the analysis of available options to turn implementation into a sector-wide, multi-actor, inclusive endeavor, through which the roles of the funder and the external process facilitator can be gradually reduced)

government or a consortium of large funders is being planned. Too often ECD funding has concentrated on either the supply side (the individuals providing evaluation services) or the demand side (the commissioners of evaluations) of the equation. However, at a time when roles often overlap such a distinction no longer appears to be reasonable.

- For organizations that are implementing ECD strategies, as well as local organizations being targeted by ECD programs, ECD program objectives or organizational visions for evaluation should fit well with their respective institutional or organizational vision.

Enhancing Evaluation Capacity: The Equity Paradox

Considering the observed gaps in the targeting of ECD programs across funding agencies in the past, and based on the results of key ECB and ECD theoretical frameworks in use, this section provides a list of suggestions on how to make ECD targeting more inclusive and effective in the future (box 6.1).

IDEAS'S CONTRIBUTION TO A PARADIGM SHIFT IN CONTEMPORARY CAPACITY DEVELOPMENT THINKING AND PRACTICE

A Radically New Perspective on Evaluation Capacity Development

The role that IDEAS has played so far, and intends to play, in the area of ECD in the future is important and timely for three main reasons. First, it allows revitalization of the discourse on ECD among IDEAS members eight years after the IDEAS Global Assembly that was organized around this theme in Johannesburg, South Africa in 2009.[5] Second, it is likely to provide guidance in the production of some concrete tools and checklists that IDEAS members could use to enhance the effectiveness of the ECD work in their respective fields in the future. Third, it would build and expand upon the work that IDEAS has already done on evaluation competencies. Overall, the work of IDEAS in this area, as envisioned in the mission of the newly created Evaluation Capacity Development International Topical Interest Group (ECD ITIG), aims to build more consensus among members of IDEAS from different regions on what it means to work with organizations and governments on evaluation capacity in a more contextualized and sustainable manner.

[5] "Getting to Results: Evaluation Capacity Building and Development. For more details, please visit the conference website: https://ideas-global.org/2009-conference/.

BOX 6.1 **Key suggestions to make ECD strategy more inclusive and effective**

1. **Understand the specific ECD ecology where you are working.** In conducting a mapping of the major institutions—both within and outside of the government sphere, including VOPEs, academia, and the private sector that have demonstrated interest in ECD in the past—the identification of individuals as well as specific units with a more vested interest in evaluation is strongly recommended, so as to avoid the personalization of the evaluation function, which would then be exposed to the risk of collapse in case of staff turnover.
2. **Identify some common nationwide goals and objectives that all ECD stakeholders could be encouraged to contribute to.** Such goals, ideally aligned with the Sustainable Development Goals, do not need to be perceived as imposed from the outside, and should be consistent to the largest extent possible with the each targeted actor's mission and objectives.
3. **Build institutional incentives from within.** The incentives that are made available, including the identification and rewarding of champions, should not focus on meeting preset performance agreements, and should promote instead the ECD actor's ability to wonder and ask questions about how to turn short-term and isolated tactics into long-term and sustainable strategies.
4. **Conduct a participatory ECD diagnostic assessment.** Through such assessment, it will be important to foster opportunities for self-reflection and mutual learning. In this vein, it would be important that such assessment focus on a selected number of the organizational and environmental factors and processes discussed in this chapter.
5. **Develop a national ECD strategy and, depending on the scope of your intervention, put into place a national ECD task force.** Far from being a logical framework or road map, a national evaluation strategy is to be regarded as a work in progress and a living document, setting general objectives and leaving ECD stakeholders space to come up with creative and innovative ways to achieve the agreed-upon objectives. This phase might require addressing some key issues, such as the creation of dedicated evaluation units with three specific responsibilities—compiling a database of evaluation data; conducting data analysis; and, fostering dissemination of evaluation findings—as well as the establishment of partnerships between different departments within the same organization. With respect to an ECD task force, it is advisable that a variety of actors with different functions and roles (from both the public and private spheres), as well as entities whose membership cuts across different spheres, be involved.

(continued)

BOX 6.1 *(continued)*

6. **Focus on strengthening the capacity of local actors (both users and providers of evaluation services), and develop an opportunistic joint exit strategy**, in close collaboration with the individuals, organizations, and institutions involved in an ECD program. Although not operationalized from the outset, an exit strategy needs to be conceptualized from the very beginning of an ECD intervention. In order to advance sustainability, it is of utmost importance to have a very strong leader in place, who is capable of innovation and available to promote internal structural changes and the reallocation of budget resources based on the organization's needs and the changing contextual opportunities (what I have called "responsible systemic-ness").
7. **Evaluate the progress of your ECD over time.** In order to measure the effectiveness of your ECD support program, you need to make sure that your target audience as well as the type of capacity development you are trying to support (according to the interests and needs expressed by in-country ECD stakeholders) have been clearly defined at the time of ECD strategy development.

The Organizational Principles Driving IDEAS's ECD Efforts

Cognizant of the peer-reviewed literature produced to date, and building upon the understanding of contemporary evaluation practices, IDEAS is striving to create an enabling environment in which an authentic evaluation culture can flourish among its members in the future. The IDEAS ECD strategy rests on six key organizational principles that have been identified as among the most influential in the development of an evaluative culture, especially at the global level:

Membership diversity. The more diverse the membership of IDEAS in terms of roles and functions, the more likely it is that IDEAS will be able to affect the national evaluation discourse within both the public and private spheres.

Decentralized leadership. The more that IDEAS leadership is shared, and the larger the availability of channels through which members can contribute to IDEAS decision-making processes, the better the compliance with internal governance rules will be.[6]

[6] According to IDEAS current organizational set-up, all of the world's regions are equally represented in the Board. Individual board members coordinate, too, with the national and regional evaluation associations falling within their respective geographical spheres of competence.

Frequent diagnostics of both IDEAS capabilities and organizational processes. The more frequently capability assessments are conducted—for example, annually—and the more promptly identified weaknesses are addressed, the more successfully IDEAS will be able to fulfill its mission.

Availability and continued monitoring of the IDEAS theory of change. The more often the IDEAS theory of change is available and is revised, based on its program development and the findings of capability assessments carried out among its members, the more often the assumptions underlying it will be monitored, and programmatic improvements will be made.

Promotion of effective international and external communication. The more well-articulated a communication strategy is, the more social and political legitimacy IDEAS ECD work could benefit from it.

Availability of opportunities to members for disseminating, exchanging, and developing ideas, theories, and concepts about evaluation. The approval of a new publication and dissemination policy by the IDEAS Board in May 2017 is an important milestone in the association's history. The process of publishing high-quality, peer-reviewed papers is an integral part of the IDEAS ECD ITIG work plan, and of that of all other actors within the association who have an interest in capacity development and professionalization.

By enabling its members to publish original work on topics related to international development evaluation, IDEAS will be able to attain the following objectives in the short term:

- Give visibility to practitioners and decision makers with no prior track record of publications
- Disseminate evaluation-related ideas on topics and/or countries rarely discussed in the mainstream peer-reviewed evaluation literature
- Encourage its members to collaborate in documenting and writing about their own evaluation practices on a more regular basis—that is, not only in response to the call for conference proposals launched every two years before the IDEAS General Assembly
- Use the publication of articles and other items posted on the IDEAS website as an entry point for further dialogue, and for mutual intellectual and professional enrichment among its members

In the medium and long term, the new policy is expected to elicit a stronger sense of personal belonging to IDEAS, and that, as a result, will assist in advancing and furthering IDEAS members' practice of development evaluation, through strengthening their capacities, and their uptake of innovative evaluation methodologies.

The Vision and Values Driving IDEAS's ECD Efforts

IDEAS makes a constant effort to promote a more just, equitable, and democratic design for the management, funding and evaluation of ECD work. By questioning paradigms that have dominated the international development discourse over the last 10 years, IDEAS has called upon evaluation practitioners and international partners to engage in evaluation and in capacity development with justice, sustainability, and transformative change in mind. In so doing, IDEAS is encouraging more reflection among members on what kinds of postures evaluation practitioners could and should have toward issues of equity, cultural competence, type and quality of evidence, sustainability, and the use of evaluation. Furthermore, this work promotes a more systematic appreciation of the cultural and linguistic diversity (what I call the "equality" of differences) within the IDEAS as well as the rest of the evaluation community. The ECD ITIG, for instance, complements the other IDEAS ITIGs, and helps to strengthen the cultural and linguistic diversity of the association's membership, as a strategy for enhancing the sustainability of future evaluation endeavors. As part of such a strategy, some of the key IDEAS resources will be translated into languages other than English (French and Spanish, among others), and stronger links with other regional evaluation conferences through the establishment of joint projects will be strengthened.

Technical Considerations Driving IDEAS's ECD Efforts

Besides the foundational and more vision-related principles inspiring its future ECD strategy, a number of technical considerations are driving IDEAS's ECD work, as follows:

- Given the lack of an effective decentralization of the M&E function in many countries, it has proven extremely difficult to promote a defused culture of evaluation through a top-down approach. Therefore, through the involvement of members at the country and subnational levels, IDEAS is attempting to extend the in-country boundaries of the evaluation culture.
- Due to the relatively weak data analysis skills among national evaluators, IDEAS aims to enhance the statistical as well as the qualitative analysis of their members through the use of webinars and other publications.
- In response to misconceptions about evaluation, or "reductive" understanding of the purposes of evaluation,[7] IDEAS will

[7] Many evaluation commissioners and planners believe that evaluation consists in verifying the compliance between what happens on the ground and the original objectives spelled out in the project logframe, with almost no reference to unexpected impact. This is what I refer to as the "RBM-ization of the evaluation function" (Tarsilla 2014b). As a result of such limited interpretation of evaluation objectives, the risk is that evaluation could become highly centralized, with little room left for conducting independent and equity-based evaluations of programs and policies.

increasingly promote critical publications, blogs, and online discussions on the IDEAS website or on LinkedIn, as a way to promote a more exhaustive definition of the evaluation function.

CONCLUSION

This chapter encourages funders as well as practitioners to rethink the way they plan, implement, and evaluate capacity development efforts. Based on a review of existing gaps in contemporary capacity-building practices in the evaluation arena, this chapter calls on all international development actors globally: to adopt a more contextually relevant, adaptive, equitable, inclusive, and democratic definition of ECD; to conduct more exhaustive diagnostics of both capacity and processes among the organizations and entities targeted by their ECD strategies; and to assess the distribution of power and resources within the systems where ECD strategies are expected to be implemented. Lastly, by documenting the current IDEAS initiatives that are aimed at maintaining and promoting evaluation capacity at several levels (national, regional, and global), this chapter attests to the association's leadership in the ECD arena.

REFERENCES

ACBF (African Capacity Building Foundation). 2016. *African Union Agenda 2063. Capacity Development Plan Framework: Buttressing Implementation of the First 10-Year Plan "The Africa We Want."* Harare, Zimbabwe: ACBF.

Argyris, C., and D. Schön. 1996. *Organizational Learning II: Theory, Method and Practice*. Reading, MA: Addison Wesley.

AU and NEPAD (African Union and New Partnership for Africa's Development Planning and Coordinating Agency). 2012. *Africa's Capacity Development Strategic Framework*. Midrand, South Africa: NEPAD.

Bamberger, M. 2009. *Institutionalizing Impact Evaluation within the Framework of a Monitoring and Evaluation System*. Washington, DC: The World Bank.

Boyle, R., D. Lemaire, and R.C. Rist. 1999. "Incentives, Constraints and Culture-Building as Instruments for the Development of Evaluation Demand." In *Building Effective Evaluation Capacity*, R. Boyle and D. Lemaire, eds., 153–77). New Brunswick, NJ: Transaction Publishing.

Bohni, Nielsen S., and K. Attström. 2011. "Perspectives on Evaluation Capacity Building." In *Influencing Change: Building Evaluation Capacity to Strengthen Governance*, Ray C. Rist, Marie-Helene Boily, and Frederic Martin, eds. Washington, DC: World Bank.

Compton, D.W., M. Baizerman, and S.H. Stockdill, eds. 2002. "The Art, Craft, and Science of Evaluation Capacity Building." *New Directions for Evaluation* 93. San Francisco: Jossey-Bass.

Fukuyama, F. 2004. *State-Building: Governance and World Order in the 21st Century*. Ithaca, NY: Cornell University Press.

Fung, A. 2003. "Recipes for Public Spheres: Eight Institutional Design Choices and Their Consequence." *Journal of Political Philosophy* 11: 338–67.

Gaventa, J., and G. Barrett. 2010. "So What Difference Does It Make? Mapping the Outcomes of Citizen Engagement." IDS Working Paper 347. Institute of Development Studies, Brighton.

GIZ (German Agency for International Cooperation). 2013. *GIZ's Monitoring and Evaluation Policy*. Bonn: GIZ.

Lennie, J. 2005. "An Evaluation Capacity-Building Process for Sustainable Community IT Initiatives: Empowering and Disempowering Impacts." *Evaluation* 11: 390-414.

Lusthaus, C., M.H. Adrien, G. Anderson, F. Carden, and G.P. Montalvan. 2002. *Organizational Assessment. A Framework for Improving Performance*. Ottawa/Washington, DC: International Research Development Centre/Inter-American Development Bank.

Mayne, J. 2008. "Building an Evaluative Culture for Effective Evaluation and Results Management." ILAC Working Paper 8. Institutional Learning and Change Initiative.

Morgan, P. 1998. *Capacity and Capacity Development: Some Strategies*. Hull: Policy Branch, CIDA.

Norad (Norwegian Agency for Development Cooperation). 2006. "Instruction for Evaluation Activities in Norway's Aid Administration."

OECD (Organisation for Economic Co-operation and Development). 2006. *The Challenges of Capacity Development: Working towards Good Practice*. Paris. OECD.

Ortiz, A., and P. Taylor. 2009. "Learning Purposefully in Capacity Development: Why, When and What to Measure?" Opinion Paper. Brighton, Sussex: Institute of Development Studies.

Pasteur, K. 2006. "Learning for Development." In *Relationships for Aid*, R. Eyben ed. London: Earthscan.

Rist, Ray C., Marie-Helene Boily, and Frederic Martin, eds. 2011.*Influencing Change: Building Evaluation Capacity to Strengthen Governance*. Washington, DC: World Bank.

Solomon, M.J., and A.M. Chowdhury. 2002. "Knowledge to Action: Evaluation for Learning in a Multi-Organisational Global Partnership." *Development in Practice* 12: 346–54.

Stevenson, J.F., P. Florin, D.S. Mills, and M. Andrade. 2002. "Building Evaluation Capacity in Human Service Organizations: A Case Study." *Evaluation and Program Planning* 25: 233–43.

Sida (Swedish International Development Cooperation Agency). 2007. *Looking Back, Moving Forward. Sida Evaluation Manual*. 2nd revised ed. Stockholm: Sida.

Tarsilla, Michele. 2012. "From Building Evaluation Capacity to Supporting Evaluation Capacity Development: The Cases of Democratic Republic of Congo, Niger, and South Africa." PhD dissertation, Western Michigan University.

———. 2014a. "Evaluation Capacity Development in Africa: Current Landscape of International Partners' Initiatives, Lessons Learned and the Way Forward." *African Evaluation Journal* 2 (1).

———. 2014b. "From RBM-ization to Normalization: A Field Practitioner's Reflection on ECD Current Trends." OECD Newsletter.

Trochim, W. 2006. "An Evaluation Culture." *Research Methods Knowledge Base* website.

Ubels, J. 2010. "Stimulating the Provision of Local Capacity Development Support." In *Capacity Development in Practice*, J. Ubels, Na-A. Acquaye-Badoo, and A. Fowler, eds. New York: Earthscan.

UNESCO (United Nations Educational, Scientific and Cultural Organization). 2015. *UNESCO Evaluation Policy*. 196 EX/24.INF. UNESCO.

USAID (United States Agency for International Development). 2011. *Evalution: Learning from Experience. USAID Evaluation Policy*. Washington, DC: USAID.

Wiesner, E. 1997. "Evaluation, Markets, and Institutions in the Reform Agenda of Developing Countries." In *Evaluation for the 21st Century: A Handbook*, E. Chelimsky and W. R. Shadish, eds., 189–200. Thousand Oaks, CA: Sage Publications.

Wijayatilake, D. 2011. "Working toward Development Results: The Case of Sri Lanka." In *Influencing Change: Building Evaluation Capacity to Strengthen Governance*, Ray C. Rist, Marie-Helene Boily, and Frederic Martin, eds., 13–22. Washington, DC: World Bank.

World Bank. 2005. *Building Effective States, Forging Engaged Societies*. Washington, DC: The World Bank.

Chapter 7

Mentoring Emerging Evaluators - Sharing Experiences from the Global South

Awuor Ponge, Taiwo Peter Adesoba, Ahmed Tammam, and Tara Devi Gurung

Abstract. *This chapter describes a mentoring program the International Development Evaluation Association (IDEAS) is implementing to support emerging evaluators. Building a business case for mentoring, the authors attempt to establish evidence for what has and has not worked, and suggest how the program can be carried out effectively. Links between mentoring and the professionalization of evaluation, and the potential benefits of the IDEAS program are discussed, as well as different models of mentoring; mentoring policies and procefdures; how they should be developed, and by whom; the importance of recognizing mentors for the work they do; and reverse mentorship, with young evaluators mentoring older professionals in social media and the use of digital technology. Summaries of presentations made by three of the authors at a panel discussion of young and emerging evaluators at the 2015 IDEAS Global Assembly describe mentoring experiences in Nigeria, the Middle East and Eurasia, and Nepal. The*

Awuor Ponge, African Policy Centre, awuorponge@gmail.com; Taiwo Peter Adesoba, East Tennessee State University College of Public Health, oluwasesan@gmail.com; Ahmed Tammam, Independent Consultant, dr.egypt85@gmail.com; Tara Devi Gurung, Nepal Red Cross Society, gurung.tara@gmail.com.

chapter offers conclusions drawn from discussions about the pilot mentoring program, and what it portends for young and emerging evaluators, especially in the Global South, as they position themselves in readiness for the evaluation of the Sustainable Development Goals.

Emerging evaluators often lack the portfolio of experience required of professionals. One solution to this challenge lies in strengthening the professional capacity and credibility of less experienced evaluators. The International Development Evaluation Association (IDEAS) is piloting a mentoring program for young and emerging evaluators to create opportunities to help them build specific professional skills.

Many young evaluators, as well as those entering the field in a career shift, are disadvantaged in terms of consultancy and job placement because of a lack of experience. Over the past decade, mentoring has proliferated as an intervention strategy for addressing the needs of young people for adult support and guidance throughout their career development. This chapter draws on experiences shared at a panel discussion at the Global Assembly of IDEAS in Bangkok in October 2015. The specific aims of this discussion were to allow young and emerging evaluators to share some of the challenges they are experiencing as they enter the evaluation profession; to brainstorm on how context-specific challenges can be addressed in such a way as to set emerging evaluators on a stable footing in the profession; and to share best practices from across the continents so that emerging evaluators can learn from the challenges of others, and how they have been addressed.

There is ample evidence of the positive contribution of mentoring to improvement in skills development, social and professional competence, and intellectual development, as well as the development of the vocational skills needed for professionalization of the evaluation discipline, while positioning mentees for professional satisfaction in their practices.

IDEAS REACHING OUT TO YOUTH

IDEAS is a global membership organization focused on the evaluation of development that brings together evaluation practitioners from all the corners of the world to help develop their professional skills while enhancing networking among members and recruiting others into the evaluation profession. In order to promote the inclusion of youth, IDEAS has initiated a youth membership category for evaluators up to age 30, with a reduced annual fee. The mentoring program is to be spearheaded by young evaluators themselves, under the guidance of senior evaluators as mentors.

ABOUT EMERGING EVALUATORS

Emerging evaluators often lack the portfolio of experience that is required by potential employers and the commissioners of consultancy assignments during the hiring process. One possible solution for enhancing the skills set and credibility of less experienced evaluators involves the establishment of

a mentoring scheme that pairs emerging evaluators with more experienced counterparts. Established evaluators are able to mentor upcoming evaluators in various ways, including working with them on assignments. Through involving youth in evaluation, we are also enhancing their individual development and encouraging their active involvement in decisions that will affect their lives. The strategy involves setting up partnerships with development organizations globally in order to create opportunities for students and youth evaluators to benefit from the advice of senior evaluators, and give them the hands-on experience they need to build specific professional skills. A strategy known as reverse mentorship will also be employed, in which youth evaluators will mentor older members in new and emerging trends in development evaluation, including but not limited to the use of digital technology and social media.

THE CONCEPT OF MENTORING

Mentoring programs for youth are commonplace in today's society: more than 5,000 such programs in the United States serve an estimated 3 million young people (MENTOR/National Mentoring Partnership 2006, as quoted in DuBois et al. 2011). In a typical program, each youth is paired with a volunteer from the community, with the aim of cultivating a relationship that will foster the young person's positive development and wellbeing (DuBois et al. 2011).

A mentoring relationship is most often defined as a professional relationship in which an experienced person (the mentor) assists a less-experienced person (the mentee) in developing specific skills and knowledge that will enhance the mentee's professional and personal growth. Evaluators are educators: their success is judged by what others can learn from their work. Mentoring is generally viewed as one component of a more comprehensive youth development strategy: these strategies can help youth gain the competencies they need to meet the challenges of adolescence and become successful adults (Foster 2001).

Over the past decade, mentoring has proliferated as an intervention strategy for addressing the needs of young people for adult support and guidance in the development of their careers. Widespread expansion of youth mentoring programs in the United States was inspired by the release of a report on an evaluation of the Big Brothers Big Sisters of America, a community-based mentoring program. Findings from this research provided evidence of associations between mentoring and a range of positive youth outcomes, and were widely embraced by policy makers and practitioners (Tierney, Grossman, and Resch 1995). Sharing experiences is a way of building community, highlighting commonalities, engaging in practices of cathartic release, and often shining light on other matters that might otherwise remain hidden. To share experiences is to articulate them, and to articulate them is to gain power over them, rather than to feel "stuck" (Vallabha 2015).

As human beings, we live in social groups where we learn our norms, values, and behaviors by the example and coaching of others. Mentoring happens in all organizations, whether it is fostered as a development strategy or encouraged as an informal process. People are constantly learning

from others, adopting modeled behaviors and attitudes, and absorbing the culture and perceived values of the organization or professional environment through their personal interactions with other members of the organization.

In the context of the IDEAS mentoring program, it is not age alone that affects the relationship between a mentor and a mentee. A mentee may be older, but changing careers. Or he or she might be an experienced evaluator who needs help learning how to use qualitative analysis software from a colleague who has experience with this skill.

DEVELOPING THE CASE FOR MENTORING: ESTABLISHING THE EVIDENCE

Because individuals may experience mentoring at various life stages, it is not surprising that there are three distinct streams of mentoring: youth mentoring; academic mentoring; and workplace mentoring. Youth mentoring involves a relationship between a caring, supportive adult and a child or adolescent (Rhodes 2002).

Mentee motivation and involvement in the evaluation profession may be greatly influenced by mentoring. Role modeling can expose mentees to the field of evaluation and to social opportunities that can open their eyes to different possibilities and motivate them to seek out new experiences. Those who have been mentored are more likely to see the need to go the extra mile in order to be fully engrossed in the profession: this can involve many different matters currently under debate regarding professionalization, including certification and accreditation, as well as other forms of recognition.

A mentoring relationship can promote career success. Mentors can impart specific knowledge and expertise that contributes to mentee learning and skill development (Eby et al. 2008; Kram 1985). Mentors can also facilitate professional networking by introducing mentees to influential individuals within academic and organizational contexts: these important career contacts can in turn lead to career success in terms of salary, promotions, and job offers (Kram 1985). Anecdotal evidence abounds of those who have been successful in evaluation practice as a result of the networking and informal mentoring they have received through working with more advanced or senior evaluators. This certainly is a positive step toward professionalizing not only the individual, but the discipline as well.

A study of youth development interventions concluded that a wide range of youth development approaches, including mentoring, result in positive behavioral changes, such as improved interpersonal skills and relationships, and increased self-control and academic achievement (Foster 2001). For example, the mentored professional will most certainly see the need to pursue academic or professional courses that are geared toward building their capacity in evaluation.

At IDEAS, the emphasis is on workplace mentoring, which occurs in an organizational setting, and the purpose of which is the personal and professional growth of the mentee (Kram 1985). The mentor may be a supervisor; someone within the organization who is outside the mentee's chain of command; or an individual in another organization (Eby 1997). Mentoring at

different developmental stages tends to serve different functions or purposes. Youth mentoring is often aimed at reducing risky behavior, or improving social and academic functioning; academic mentoring tends to target student retention, academic performance, and adjustment to college life; while workplace mentoring aims to enhance the personal and career development of employees (Eby et al. 2008).

A study of the top 200 executives in 50 large U.S. companies revealed that mentoring is one of the most effective tools in personnel development (Chambers et al. 1998). Another study, by the Institute of Management, involving 1,500 U.K. managers revealed that mentoring is one of the two most powerful leadership development tools used in organizations. It has also been reported that 71 percent of Fortune 500 and private companies use mentoring in their organizations, and that 77 percent of U.S. companies surveyed said that mentoring had improved both the retention and performance of employees.[1]

In addition to correlations with higher educational aspirations, the research shows that mentoring is correlated with other positive developmental outcomes, including changing attitudes (higher self-esteem, and stronger relationships with adults, including with teachers and peers) and better behavior (Bruce and Bridgeland 2014). In 2002, DuBois and colleagues published a meta-analytic synthesis of findings from 55 evaluations of youth mentoring programs that had been published through 1998 (DuBois et al. 2002). These findings indicated that, on average, youth participating in mentoring programs had benefited significantly in each of five outcome domains: emotional/psychological, problem/high-risk behavior, social competence, academic/educational, and career/employment (DuBois et al. 2002, 2011).

Studies have shown that many young adults are entering the labor force without even the limited skills that are necessary to attain a job in the first place, such as interviewing skills, conflict resolution, and effective communication (Eccles and Gootman 2002). Mentoring has been linked with a myriad of intellectual skills and development, including good decision-making skills, in-depth knowledge of more than one culture, knowledge of both essential life skills and vocational skills, and rational habits of mind, such as critical thinking and reasoning skills. One study found that nearly all young adults who had formal mentoring relationships (95 percent) found these experiences to be "helpful," half of which (51 percent) found the relationship to be "very helpful." Similarly, nearly all youth in informal mentoring relationships (99 percent) said their experience was "helpful," seven in 10 (69 percent) reporting it as "very helpful" (Bruce and Bridgeland 2014).

All of these studies can be summed up in one statement: there is ample evidence of the positive contribution of mentoring to improvement in skills development, social and professional competence, and intellectual development, as well as the development of the vocational skills that are needed for professionalization of the discipline: it also positions mentees for professional satisfaction in the practice.

[1] Chronus, https://chronus.com/how-to-use-mentoring-in-your-workplace.

WHAT IS THE ADDED VALUE OF A MENTORING PROGRAM FOR IDEAS?

Organizations recognize that workforce demographics have changed dramatically in recent years. Youth are becoming employed in large numbers, thanks to the global youth bulge. More and more graduates are joining the workforce as junior professionals in evaluation practice, while some are developing interest in the profession while they are still in university. There is a need to provide sufficient growth opportunities for potential future professionals in the practice. Senior evaluators also need to be mentored in new and emerging evaluation methodologies, including the use of digital technology and social media: they could benefit from the proposed reverse-mentoring program.

These are the potential benefits of the IDEAS mentoring program for the organization:

- Recruitment of new members to the organization and discovery of talent
- Development of leadership for the future survival and prosperity of IDEAS
- Communication of values, goals, and plans of the organization globally
- Demonstration of personal and professional standards among members
- Implementation of equity initiatives
- Fostering of shared values and teamwork
- Building a strong global learning organization in evaluation practice
- Development of cross-organizational networks
- Increase in morale and motivation among both junior and senior professionals

DEVELOPMENT OF A MENTORING PROGRAM

The IDEAS Board has appointed a mentoring program coordinator, who proposed the program. The program coordinator then formed a task force of five people, which has been approved by the Board to develop the mentoring program policy, as well as to oversee its implementation. Members of the task force represent a cross-section of the organization, including potential mentors and mentees, as well as stakeholders who bring value to the process. The duties of the task force include the following:

- Determining the goals of the program
- Choosing the proper mentoring model
- Developing a memorandum of understanding (MOU) template for the mentoring partnership
- Negotiating the MOU with major institutions commissioning development evaluations
- Selecting eligible evaluation training programs/institutions
- Defining criteria for mentors and mentees

- Defining other critical components of the program
- Matching the participants
- Monitoring the pilot
- Evaluating the results at the end of the pilot program

The mentoring panel at the Global Assembly was sponsored by the U.K. Department of International Development (DFID), among others, but no other funds could be raised. Thus, many of the more ambitious plans turned out to be unrealistic. Similar initiatives in the European Evaluation Society and the American Evaluation Association were largely voluntary in nature; this seems to be a more realistic way forward for IDEAS as well.

PURPOSE AND OBJECTIVES OF THE IDEAS MENTORING PROGRAM

The mentoring program at IDEAS aims to achieve the following:

- **Induction and skills enhancement.** Helps new members settle into the organization; facilitates potential and skill development for new members, both young evaluators and those just entering a career in evaluation; seeks to enable skills to be passed on by experienced, highly competent professionals to others who need to develop specific skills.
- **Networking and career development.** Helps both mentors and mentees in the planning, development, and management of their careers; helps them become more resilient in times of change; and more self-reliant in their careers. Offers young professionals visibility and the opportunity for networking, which helps them to explore and plan their career pathways. Also helps both mentors and mentees gain a greater awareness of opportunities and activities that can broaden their professional experience.
- **Education support and practical orientation.** Helps bridge the gap between theory and practice; complements formal education and training through sharing the knowledge and hands-on experience of competent practitioners. Offers mentees the opportunity to acquire new knowledge and skills by observing and understanding the mentor's practical experience.
- **Leadership and development of competencies.** Encourages the development of leadership and professional competencies that are more easily gained through example, guided practice, or experience than through theoretical education and training.
- **Global visibility and organizational development.** Will expand IDEA's culture of cooperation and commitment through sharing the values, vision, and mission of the organization, and will give IDEAS an enhanced visibility globally.

The program proposes the following models of mentoring:

- **One-to-one mentoring**—matches one mentor with one mentee
- **Group mentoring**—assigns one mentor to work with several mentees
- **Team mentoring**—involves more than one person working with the same mentee
- **Computer online mentoring**—uses computer-based opportunities to develop relationships through online communication
- **Peer mentoring**—young people experienced in evaluation mentor other young people who are just entering the evaluation profession

In order to realize and sustain the program, IDEAS needs to find mechanisms for arranging mentoring on a voluntary basis, without giving up on the effort to mobilize resources for the program that are needed for purchasing mentoring software; covering the administrative costs for staff managing the program; and the costs of training for mentors and mentees, among other things. There will also be a need to provide mentors with formal recognition for all they do for young and emerging evaluators. This recognition can take many forms, such as awards—for example designating a mentor/mentee match of the year—as well as gifts and/or letters of appreciation.

Three of the case studies presented at the Global Assembly in Bangkok in October 2015 are summarized below.

NIGERIA CASE STUDY

In a presentation titled "Evaluation Capacity Development for Emerging Evaluators: A Nigerian Experience," Taiwo Peter Adesoba observed that describing an emerging evaluator is sometimes a tricky task, because there seems to be no globally agreed-upon definition in terms of age, educational requirements, job experience, and so on. Just as a plant has hurdles to overcome when emerging from the soil, so emerging evaluators have particular needs while they are trying to establish themselves in the evaluation profession. A major challenge for young evaluators, especially in low- and middle-income countries, is weak evaluation capacity, which leads to their exclusion from evaluation activities. With the growing number of youth-led organizations, especially in Africa and Asia, more attention is being given to the monitoring and evaluation (M&E) components of their projects. Young development professionals are keenly interested in evaluating the impact of their own work: the demand for accountability from policy makers and other project implementers in their communities is also a factor. Although young evaluators are passionate about implementing evaluation assignments, the technical skills required to properly fulfill this passion is poor. This makes evaluation capacity development (ECD) a necessity for young and emerging evaluators.

The ECD Project in Nigeria is targeted at young evaluators aged 18-30 with less than two years' experience in M&E. The project was conceptualized following a survey among young evaluators in which about 81 percent of respondents said they had never had formal training in M&E: 92 percent of them did not belong to any voluntary organizations for professional evaluation (VOPEs), and 100 percent of them said they needed mentoring. The

objectives of ECD programs are to improve youth inclusion not only in project implementation, but in its evaluation; to promote and advance the practice of the evaluation profession; and to increase youth participation in decision making for sustainable development. The participants in the first phase of the ECD were 24 (7 males; 17 females) emerging evaluators ranging from 21 to 33 years who work for local nongovernmental organizations (NGOs) in the Ondo and Ekiti states. They take on M&E roles in their organizations, but have only limited experience. About 20 percent of the trainees had attended M&E training previously. None of them belonged to any VOPEs in Nigeria, and many had never attended an evaluation conference. The ECD covered the basics of M&E, and its application to their organizations' projects.

Some of the lessons learned during the two-day training were tailored toward the project areas of the trainees. Assignments and group work were aimed at the areas of HIV prevention, economic empowerment, and gender programs. After two days of intensive training, group work, and discussions, the participants showed a good understanding of the basics of M&E. Some have gone on to find online evaluation webinars, and two of the trainees have gotten internship opportunities with NGOs in other parts of the country. However, better results could be achieved if the trained young evaluators had access to mentors who could provide them with additional guidance.

Adesoba concluded that capacity development of emerging evaluators for sustainable development will bring about better results if well-qualified evaluation professionals are available to provide on-the-job mentoring, through either short-term internships or evaluation job placements.

THE MIDDLE EAST AND EURASIA: SAVE THE CHILDREN CASE STUDY

In his presentation "The Future of Evaluation," Ahmed Tammam discussed the challenges faced by emerging evaluators in the Middle East and Eurasia (MEE) region. He observed that M&E is a critical component in the advancement of evidence-based interventions, and that it consequently enhances efficiency, transparency, and accountability. Evaluation is a profession, and grooming new leadership for the future of this profession is needed: thus, investment in developing the capacities of young evaluators is an emerging priority.

Tammam noted that the evaluation profession is far ranging. Therefore, support from experienced evaluators is key in order for young evaluators to advance in such a robust profession, which has so many specific and different areas of work. Evaluation is a growing career path, but visibility, employment, and networking opportunities can be very limited for young evaluators.

He further noted that talented young evaluators in the MEE region are increasingly lobbying donors to fund only successes, and described a case study of Save the Children in the MEE region as an example. Save the Children is a child rights–based organization, working in 12 countries in the Middle East and Eurasia, with operations that support children in developmental contexts (as in Egypt and Georgia), as well as in an emergency contexts (as in Syria, Ukraine, and Yemen). With the expansion of needs in the

region and the limitation of resources, there is a push to focus only on what works best for children. Consequently, there is a need for effective evaluation, and thus for evaluators.

In 2014, the program quality department of Save the Children at the MEE regional office began a process of investing in the talents of local emerging evaluators by recruiting an evaluator from a country office to manage the M&E tasks on the regional level. One of the mandates of the recruited evaluator was to begin to establish a cadre of talented young evaluators. A capacity-building process for emerging young evaluators was undertaken, promoting intra-learning and providing opportunities for these individuals to gain more knowledge and enhance their experience as emerging evaluators. This interactive learning program was mentored by senior program-quality staff who helped the emerging evaluators map the real needs in their countries.

As a result, an interactive mapping exercise of local M&E needs and capacities was conducted by the emerging evaluators, and a regional think tank was created. Through this process all of the young evaluators gained on-the-job experience. Two of them offered to support other country offices; another five were given the authority to be part of the operationalization process of the Vision and Position Paper of the Eurasia region (Albania, Armenia, Georgia, Kosovo, and the Northwest Balkans), which aims to align the efforts of five countries to work together on subregional programs in order to maximize benefits for children and to efficiently utilize the limited funding opportunities, by pushing donors to fund success.

In conclusion, Tammam noted that strengthening young and emerging evaluators through enhancing the learning environment and providing opportunities for mentorship are key not only to developing the young evaluators' future, but the future of the evaluation profession as well.

NEPAL CASE STUDY

Tara Devi Gurung presented a case based on her anecdotal experience as an emerging evaluator in Nepal, and also from a desk review of the available literature about development evaluation, and the role of the young and emerging evaluators in Nepal.

In her presentation, "Evaluation Challenges and Opportunity for Emerging Evaluators: A Nepalese Experience/Case," Gurung noted that evaluation, which assesses the results of policies, programs, and projects, is an integral part of the development process. Evaluation is particularly critical in the context of Nepal, which has a complex social structure, a high poverty rate, gender discrimination, dynamic forces of globalization sweeping traditional societies, and numerous development projects aimed at the population. These factors, in addition to greater competition for limited resources available for international development are pushing donors, program participants, and evaluators to seek more rigorous, but still flexible, systems for monitoring and evaluating development and humanitarian interventions.

Many of the current approaches to evaluation are unable to address the changing structure of development assistance, and the increasingly complex environment in which it operates. Innovative evaluation approaches

and practices are particularly important in such situations. However, it is challenging for emerging evaluators to carry out innovative evaluation, since this calls for a high level of expertise.

Emerging evaluators are those who have recently joined the profession and have limited experience. They are not necessarily young, although most of them are relatively young. Many of them are disadvantaged in terms of consultancy and job placement because they lack sufficient experience. There are only limited forums for sharing and disseminating best practices; developing quality and ethical protocols; enhancing the capacity of new evaluators; and researching evaluations. There are limited resources for evaluation that collects lessons drawn from evaluations around the world; produces knowledge through research undertakings; and supports the development of curricula for and carries out basic and advanced training in evaluation. Moreover, there is no academic institution in Nepal that offers a university degree in evaluation. Lack of specific acts, rules, and regulations for evaluation has posed even more challenges for evaluators in Nepal. Available guidelines focus only on governmental M&E, and do not cover other sectors.

In this context, it is difficult to raise funds for evaluation research, capacity building, and activities related to the promotion of evaluation in Nepal. The funding agencies have a tendency to support already established organizations and firms rather than new evaluators. Often the potential for innovation, and the expertise of emerging evaluators, are overlooked by the commissioners of evaluations. Better representation, and the active engagement of young and new evaluators in the decision-making process are needed in order to bring their ideas and perspectives into evaluation.

Gurung concluded that effective evaluation is crucial in order to assess the progress and impact of the efforts of the government policy to “Build a New Nepal” through accelerated development inclusive of all castes, ethnic groups, and genders. Evaluation is gradually becoming an integral part of development plans, projects, and emergency operations in Nepal. Gurung predicts that this will lead to a rise in skilled evaluators, including emerging evaluators in the near future.

CONCLUSION

The IDEAS mentoring program aims to enhance induction and skills development, networking and career development, education support and practical orientation, leadership and competencies development, global visibility, and organizational development for young and emerging evaluators.

Mentoring at a variety of developmental stages tends to serve different functions or purposes. Workplace mentoring is aimed at enhancing employees’ personal and career development: this is the kind of mentoring that IDEAS is piloting. In general, mentoring has been linked with a myriad of intellectual skills and development, including good decision-making skills, in-depth knowledge of more than one culture, knowledge of both essential life skills and vocational skills, and rational habits of mind such as critical thinking and reasoning skills. There is ample evidence of the positive contribution of mentoring to improvement in skills development, social and professional

competence, and intellectual development, as well as the development of the vocational skills that are needed for professionalization of the evaluation discipline, while also positioning mentees for professional satisfaction in their practices.

Studies have shown that, among other benefits, youth participating in mentoring programs have benefited significantly in social competence, academic and educational progression, as well as career or employment advancement. The three cases discussed in this chapter demonstrate that there is a country-level need for supporting young and emerging evaluators through mentoring in order to advance their professional competence and career prospects. These cases demonstrate further that capacity development of emerging evaluators will ensure that more well-qualified evaluation professionals are available to provide on-the-job mentoring through short-term internships or evaluation job placements. They also demonstrate that strengthening young and emerging evaluators through enhancing the learning environment, and providing opportunities for mentorship are key, not only to developing the future of young evaluators, but for the future of evaluation as a profession in general. Overall therefore, these studies build a strong business case for the mentoring of young and emerging evaluators.

ACKNOWLEDGMENTS

Special thanks to the members of the IDEAS Executive Board namely: Rob D. van den Berg (President); Susan Daet Tamondong (Vice-President), and Cristina Magro (Secretary), who have all supported this initiative since inception. We also wish to acknowledge the valuable comments received from Taiwo Peter Adesoba, Kruno Karlovcec, Jérôme Gandin, and Halcyon Louis, who are all members of the IDEAS Mentoring Policy Development Team which is leading the piloting of this initiative at IDEAS. Lastly, thanks to the panelists at the Mentoring Panel Session in Bangkok, Thailand: Awuor Ponge, Taiwo Peter Adesoba, Ahmed Tammam, Tara Devi Gurung, and Ana Borges.

REFERENCES

Bruce, M., and J. Bridgeland. 2014. "The Mentoring Effect: Young People's Perspectives on the Outcomes and Availability of Mentoring." A report for MENTOR: The National Mentoring Partnership.

Chambers, E.G., M. Foulon, H. Handfield-Jones, S.M. Hankin, and E.G. Michaels III. 1998. "The War for Talent." *McKinsey Quarterly* 3: 44–57.

DuBois, D.L., B.E. Holloway, J.C. Valentine, and H. Cooper. 2002. "Effectiveness of Mentoring Programs: A Meta-Analytical Review." *American Journal of Community Psychology* 30, 157–97.

DuBois, D.L., N. Portillo, J.E. Rhodes, N. Silverthorn, and J.C. Valentine. 2011. "How Effective Are Mentoring Programs for Youth? A Systematic Assessment of the Evidence." *Psychological Science in the Public Interest* 12 (2):57–91.

Eby, L.T. 1997. "Alternative Forms of Mentoring in Changing Organizational Environments: A Conceptual Extension of the Mentoring Literature. *Journal of Vocational Behavior* 51: 125–44.

Eby, L.T., T.D. Allen, S.C. Evans, T. Ng, and D. DuBois. 2008. "Does Mentoring Matter? A Multidisciplinary Meta-Analysis Comparing Mentored and Non-Mentored Individuals." *Journal of Vocational Behavior* 72 (2): 254–67.

Eccles, J., and J.A. Gootman, eds. 2002. *Community Programs to Promote Youth Development*. Washington, DC: National Academy Press.

Foster, L. 2001. "Effectiveness of Mentor Programs: Review of the Literature from 1995 to 2000." CRB-01-004. California Research Bureau, Sacramento, CA.

Kram, K.E. 1985. *Mentoring at Work*. Glenview, IL: Scott Foresman.

MENTOR/National Mentoring Partnership. 2006. "Mentoring in America 2005: A Snapshot of the Current State of Mentoring." Alexandria, VA: MENTOR/National Mentoring Partnership.

Merrick, L. n.d. "How Coaching and Mentoring Can Drive Success in Your Organization." White Paper, Chronus Mentoring and Talent Development Solutions.

Rhodes, J.E. 2002. *Stand by Me: The Risks and Rewards of Mentoring Today's Youth.* Cambridge, MA: Harvard University Press.

Tierney, J.P., J.B. Grossman, and N.L. Resch. 1995. *Making a Difference. An Impact Study of Big Brothers Big Sisters.* Philadelphia: Public/Private Ventures.

Vallabha, Bharath. 2015. "The Value of Sharing Experiences." *The Rough Ground* blog.