
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

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 inequality, 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, sub-national, 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 multi-dimensional 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

⁶Personal 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

⁸For 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

⁹Much 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 foundations; 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 foundations; 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 professional practice	Managing the evaluation	Professional practice: capacity to manage and deliver evaluations	Management domain
Inclusive/communicates effectively and respectfully/negotiation	Interpersonal practice		Managing the evaluation; 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-Saddik, 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. Ghore, 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.