


The role of evaluation in achieving the SDGs

Aki Yonehara¹  · Osamu Saito² · Kaoru Hayashi³ · Masafumi Nagao² · Ryu Yanagisawa⁴ · Kanako Matsuyama²

Received: 31 July 2017 / Accepted: 23 August 2017 / Published online: 7 September 2017
© Springer Japan KK 2017

Abstract This paper proposes a framework on Sustainable Development Goal (SDG) evaluation, arguing that attainment of the 17 goals and 169 related targets depends significantly on practice-based monitoring and evaluation. The SDGs' 15-year time frame can helpfully be divided into three 5-year phases: a planning phase driven by proactive evaluation and evaluability assessment, an improvement phase characterized by formative evaluation and monitoring, and a completion phase involving outcome and impact evaluations. Under these phases, in order not to miss the SDGs' fundamental philosophy of “no one left behind,” local relevance must be considered when evaluating SDG programs, particularly to capture the overarching concepts applicable across the 17 goals, such as educational dynamics and resilience.

Keywords SDG evaluation · Proactive evaluation · Evaluability · Evaluation of educational dynamics · Evaluation of resilience

Shifts in evaluation from MDG to SDG

The Sustainable Development Goals (SDGs) adopted in the UN General Assembly resolution “Transforming our World: The agenda for Sustainable Development” in September 2015 represent an inter-governmental agreement among the member states of the United Nations to promote global sustainable development based on universal principles of international cooperation and backed by their national commitment. However, the resolution is not a legally binding instrument. As such, its implementation, especially with regard to achieving the 17 adopted goals and their 169 related targets, depends significantly on the practice of evidence-based monitoring and evaluation at the national and international levels. The resolution's language expresses the underlying approach as follows:

We commit to engaging in systematic follow-up and review of the implementation of this agenda over the next 15 years. A robust, voluntary, effective, participatory, transparent and integrated follow-up and review framework will make a vital contribution to implementation and will help countries to maximize and track progress in implementing this agenda to ensure that no one is left behind (UN General Assembly 2015).

Encouraged by the relative success of quantitative tracking of the earlier Millennium Development Goals (MDGs), the governments and international organizations have embarked on the elaboration of a global SDG indicator framework, identifying more than 230 indicators thus far that correspond to the 169 targets.¹ Their initial concern

Handled by Osamu Saito, United Nations University Institute for the Advanced Study of Sustainability, Japan.

✉ Aki Yonehara
yonehara@toyo.jp

¹ Toyo University, Hakusan, Bunkyo, Tokyo, Japan

² United Nations University Institute for the Advanced Study of Sustainability (UNU-IAS), Shibuya, Tokyo, Japan

³ Faculty of International Studies, Bunkyo University, Chigasaki, Kanagawa, Japan

⁴ Akita Age Lab, Yuwatsubakigawa, Akita, Japan

¹ The total number of indicators identified was 232 as of March 2017. *Report of the Inter-Agency and Expert Group on SDG Indicators* (E/CN.3/2017/2), Annex III.

Table 1 Three 5-year phases for SDG implementation and evaluation

Phases	Activities	Evaluation concern
Phase 1 (2016–2020)	Planning and initiation of major programs	Proactive evaluation Evaluability assessment
Phase 2 (2021–2025)	Project continuation, modification, improvement, addition	Monitoring Formative evaluation
Phase 3 (2026–2030)	Project completion Follow-up	Outcome evaluation Impact evaluation

is how to monitor or measure the progress of SDG implementation. This concern is quite understandable given the enormous tasks of compiling huge volumes of data, verifying reliability and comparability, and coordinating all the efforts involved. However, the implementation review should go beyond simple measurement and should have a wider scope, also assessing whether the progress made is “equitable, relevant, and sustainable” with no one being left behind (Schwandt et al. 2016). For this reason, it is meaningful to consider the role of evaluation, not limited to the monitoring function, in enabling achievement of the SDGs.

Global SDG evaluation initiatives

Efforts to pursue broader SDG evaluation work have already started, although they are still on a modest scale. At the international level, the UN Evaluation Group (<http://www.uneval.org/>), composed of evaluation experts from various UN organizations, organized a seminar in April 2016 to address such themes as evaluability of the SDGs; evaluation for equity, equality, and non-discrimination; strengthening of national evaluation capacity; and a human security agenda for the SDGs (UN Evaluation Group 2016). Another initiative, EvalPartners (<http://www.evalpartners.org/>), started by the International Organization for Cooperation in Evaluation (IOCE), UNICEF, and evaluation experts from some international NGOs, has set up an Eval-SDGs website to address SDG evaluation issues from the point of view of civil society. A report published by leading members of this group has provided an in-depth analysis of how to incorporate the value premise of “no one left behind” in SDG evaluation (Bamberger et al. 2016). Its authors warn that evaluation methods designed to focus on the outcomes and impact of SDG interventions, especially at the national level, may cause relative neglect of process-related or contextual issues, leading to implementation failure. Yet another perspective has been offered by internationally known evaluator Patton (2015), who emphasizes that SDG evaluation should treat the entire earth as a unit of analysis.

Practical conceptualization of SDG evaluation concerns

When we consider the role of evaluation in supporting SDG achievement at the national or international level, one practical question that arises is how to cope with the long-term nature of the 2030 agenda. In other words, how should one approach evaluating a 15-year undertaking? What can one suggest to a government evaluation office charged with the task of evaluating SDGs at the national level? One practical way may be to break down the 15-year period into three 5-year phases and conceptualize the corresponding developmental issues and accompanying evaluation concerns (Table 1).

In Phase 1, SDG programs and projects are planned and initiated in all countries, and a large amount of funding is invested. In Phase 2, leading programs and projects should enter a cruising mode; some will be completed, others may be reorganized for improvement, and new programs may be started. In Phase 3, as the endpoint of the SDG time frame nears, many programs and projects are terminated and plans for follow-up and a new agenda may be taken up.

Evaluation concerns should also evolve in correspondence with the progression of phases. Phase 1 calls for proactive evaluation to ensure that the new programs and projects being initiated are set up properly and consistently with the SDGs, or to complete an evaluability assessment of some initiated programs and projects facing initial difficulties.² Phase 2 will involve monitoring evaluation, and some programs and projects nearing completion may be ready for outcome evaluation. But the agenda is for 15 years, so some projects may be reformulated and revitalized with formative evaluation. Phase 3 will require collection of final data for various indicators to facilitate outcome and impact evaluation and judgments as to whether the desired results have been attained.

² (Owen 2006, Chap. 9–10) emphasizes the important role of evaluation in the formative stages of new interventions using the notion of “proactive” and “clarificative” forms of evaluation.

Evaluation of educational dynamics

Goal 4 of the SDGs reflects an educational aim including seven specific targets, and education or human development is one of the overarching goals across multiple SDG challenges. The SDGs' educational goal has a substantively different characteristic from that of the previous global development goals, the MDGs.³ A fundamental concept of MDG performance measurement entailed establishing measurable goals, such as levels of school enrollment, in advance to be achieved during a certain period. On the other hand, the fundamental idea underlying the SDGs is more formative, and therefore it is highly difficult to set relevant indicators in advance.⁴ For example, Target 4.7 of Goal 4 states the following:

By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development.

This goal is not of the type that can be completely achieved within any designated term, and it is quite difficult to measure such an abstract target with a universally comparable indicator. To evaluate this kind of dynamic goal, it is necessary to consider *evaluability* before discussing *measurability*. Evaluation activities in the field of education often fail due to the ambiguity of purposes and plans, an ambiguity that is rooted in the dynamic and uncertain nature of education. In other words, clear, specific, and locally relevant purposes and plans may increase evaluability, and evaluable activities can receive substantive benefits from evaluation so as to improve the programs (Rossi et al. 2004). The SDG educational goal requires us to pay more attention to discussing evaluability, not only to developing new indicators.

How can discussion of evaluability be embodied in the practical process? According to program evaluation theory (Rossi et al. 2004; Smith 1990), evaluation activities should be incorporated into the whole plan-do-check-act

(PDCA) cycle, not limited to the “C” stage. Program evaluation includes five types of evaluation covering each stage of PDCA: *needs evaluation* and *theory evaluation* at the planning stage; *process evaluation* at the implementation stage; and *impact evaluation* and *efficiency evaluation* at the check stage, with the results of these evaluation activities contributing toward improving the program at the action stage. Among these five types of evaluation, *theory evaluation* plays a fundamental role to build a structure of the program by discussing among the stakeholders on the logical relationship between purpose and means. This kind of evaluation activity, a prospective evaluation at the planning stage rather than a retrospective one at the check stage, should be more relevant to the evaluation of the SDG educational goal.

Evaluation of resilience

Resilience is one of the cross-cutting and emphasized concepts frequently used throughout the SDGs. The term originated from the technical area of mechanical and engineering sciences, where it describes the properties of materials and their ability to withstand severe conditions (Hollnagel et al. 2006). The Intergovernmental Panel on Climate Change (IPCC 2014, p. 5) defined it as “the capacity of ... systems to cope with a hazardous event or trend or disturbance, responding or reorganizing in ways that maintain their essential function, identity, and structure, while also maintaining the capacity for adaptation, learning, and transformation.”

Measuring resilience in socio-ecological systems is associated with the system's abilities of reorganization, learning, and adaptation (Carpenter et al. 2001; Walker et al. 2002). Social and ecological aspects of resilience should be captured by empirical indicators, such as institutional structures, diversity of income sources, migration, and mobility, which may be affected by environmental variability such as extreme events (Adger 2000; Antwi et al. 2014). For example, Adger et al. (2005) highlighted the socio-ecological resilience of coastal areas, and Antwi et al. (2014) proposed a set of community-based resilience indicators covering three dimensions (ecological, engineering, and socioeconomic resilience) in the context of northern Ghana. Moreover, the United Nations University developed a set of 20 indicators covering five main areas: (1) landscape/seascape diversity and ecosystem protection; (2) biodiversity; (3) knowledge and innovation; (4) governance and social equity; (5) livelihoods and well-being (UNU-IAS 2014).

In spite of the existence of many conceptual assessment frameworks and models (e.g., Folke 2006; Walker and Salt 2006), localization, downscaling, and customization of

³ Goal 2 of the MDGs was an educational goal, to achieve universal primary education, with a set of indicators such as primary school enrolment and adult literacy.

⁴ (UNDESA 2016, pp 26–27) categorizes SDG indicators into three groups (Tiers 1–3). Tier 1 is the group of the indicators with an established methodology and data already widely available. Tier 2 is that with an established methodology but insufficient data coverage. And Tier 3 is that for which a methodology is being developed. Three out of five Target 4.7-indicators are categorized into Tier 3 (UNESCO-UIS 2016, p 54), and there are still 40% of all indicators categorized into Tier 3.

resilience assessment to suit a particular local context is one of the key challenges for SDG evaluation, because there are often gaps or mismatches between theory-driven indicators and local conditions. In addition to resilience indicators, we also need to identify a set of key principles about how to operationalize the assessment process, including pre- and post-assessment phases such as a follow-up phase to improve resilience based on assessment results. Two studies have explored and examined such principles for building resilience (Biggs et al. 2015; Saito et al. 2017). Building and maintaining resilience requires a sustained and continuous development, monitoring, and evaluation of intervention strategies, while engaging and collaborating with different stakeholders.

Evaluating so that no one is left behind

The ultimate test for the achievement of SDGs is whether anyone is left behind in the pursuit of the 2030 agenda (Bamberger et al. 2016). This test should not wait for the closing years of the 15-year period, but should be practiced in the implementation process of projects, programs, and policies for achieving the SDGs. Particular importance is attached to their evaluation in the initial years when massive efforts are made to launch SDG-oriented activities quickly, sometimes without due care for those who may be left behind by such efforts. Another caution to be heeded is that the current focus on elaboration of an SDG indicator system may lead to a monitoring practice that pays attention only to the quantitative information generated by the system, therefore, overlooking the intractable dark corners where someone may be left behind. Finally, given the rapid pace of technological progress and diffusion in today's globalizing world, much change of unimaginable proportions may occur in the course of the 15-year period, which may eventually alter the scope of development efforts. For example, in a rural, non-electrified village in Tanzania, a small firm called Digital Grid lends solar panels to local kiosk owners, who generate electricity, store it with the firm, and buy back the needed quantity to provide recharging services to inhabitants for their cell phones and lanterns (Japan International Cooperation Agency 2017). This type of locally based initiative can greatly reshape a community's needs and opportunities. Accordingly, evaluators of efforts toward achieving the SDGs should always keep their eyes fixed on the ground.

Acknowledgements The authors would like to express thanks for the financial support received from the Japan Society for the Promotion of Science through the Grant-in-Aid for Challenging Exploratory Research (No. 16K13348).

References

- Adger WN (2000) Social and ecological resilience: are they related?. *Prog Hum Geogr* 24:347–364
- Adger NW, Hughes TP, Folke C, Carpenter SR, Rockström J (2005) Social-Ecological Resilience to Coastal Disasters. *Science* 309:1036–1039
- Antwi EK, Otsuki K, Saito O, Obeng FK et al (2014) Developing a Community-Based Resilience Assessment Model with reference to Northern Ghana. *J Integr Disaster Risk Manag* 4:73–92
- Bamberger M, Segone M, Tateossian F (2016) Evaluating the Sustainable Development Goals: with a “no one left behind” lens through equity-focused and gender-responsive evaluations. https://www.evalpartners.org/sites/default/files/documents/eval_gender/Eval-SDGs-WEB.pdf. Accessed 19 July 2017
- Biggs R, Schluter M, Schoon ML (2015) Principles for building resilience sustaining ecosystem services in social-ecological systems. Cambridge University Press, New York
- Carpenter SR, Walker BH, Anderies JM, Abel N (2001) From metaphor to measurement: resilience of what to what? *Ecosystems* 4:765–781
- Folke C (2006) Resilience: the emergence of a perspective for social-ecological systems analyses. *Glob Environ Chang* 16:253–267
- Hollnagel E, Pariès J, Woods DD, Leveson N (eds) (2006) Resilience engineering: concepts and receipts. Ashgate Publishing, Aldershot
- Intergovernmental Panel on Climate Change (IPCC) (2014) Summary for policymakers. In: Climate change 2014: impacts, adaptation, and vulnerability. Part A: global and sectoral aspects, Contribution of Working Group II to the Fifth Assessment Report. Cambridge University Press, Cambridge, UK
- Japan International Cooperation Agency (2017) Sell electricity by measure in a non-electrified community, *mundi* 46:16
- Owen J (2006) Program evaluation: form and approaches, 3rd edn. Allen & Unwin, Crowds Nest
- Patton MQ (2015) A transnational global systems perspective: in search of the blue marble evaluation. *Can J Progr Eval* 30(3):374–390
- Rossi PH et al (2004) Evaluation: a systematic approach, 7th edn. Sage Publication, CA
- Saito O, Bofofo YA, Jasaw GS, Antwi EK, Kikuko S, Kranjac-Berisavljevic G, Yeboah R, Obeng F, Gyasi E, Takeuchi K (2017) The Ghana model for resilience enhancement in semi-arid Ghana: conceptualization and social implementation. In: Saito O, Kranjac-Berisavljevic G, Takeuchi K, Gyasi E (eds) Strategies for building resilience against climate and ecosystem changes in sub-Saharan Africa. Springer, NY
- Schwandt T, Ofir Z, Lucks D, El-Saddick K, D'Errico S (2016) The International Institute for Environment and Development (IIED) Briefing. (<http://pubs.iied.org/17357IIED>). Accessed 26 April 2017
- Smith M (1990) Program evaluation in the human services. Springer, NY
- UN General Assembly resolution (2015) 70/1 on “transforming our world: the 2030 agenda for Sustainable Development” adopted on 25 September 2015, paragraph 72
- UN Department of Economic and Social Affairs (UNDESA) (2016) Report of the secretary-general: progress towards the Sustainable Development Goals. United Nations, NY
- UN Evaluation Group (2016) Report of the United Nations Evaluation Group evaluation practice exchange 2016 seminar 25–26 April, 2016. WIPO, Geneva
- UNESCO-UIS (2016) Laying the foundation to measure sustainable development goal 4. UNESCO-UIS, Quebec

Walker BH, Salt D (2006) Resilience thinking: sustaining ecosystems and people in a changing world. Island Press, Washington, DC

Walker BH, Carpenter SR, Anderies J, Abel N, Cumming G, Janssen M, Lebel L, Norberg J, Peterson GD, Pritchard R (2002)

Resilience management in social-ecological systems: a working hypothesis for a participatory approach. *Conserv Ecol* 6(1):14