CGIAR

Food and Agriculture Organization of the United Nations

IFAD
Investing in rural people

WFP
World Food Programme
wfp.org

Enhancing the evaluability of Sustainable Development Goal 2: "End hunger, achieve food security and improved nutrition and promote sustainable agriculture"

International technical seminar

17-18 November 2015

THEME PAPERS

2015
International
Year of
Evaluation

Theme 1: The relevance of 'new metrics' for the evaluation of SDG2 – Data revolution and innovative approaches for assessing human wellbeing

Chair: John Hoddinott, H.E. Babcock Professor of Food and Nutrition Economics and Policy, Cornell University

Speakers: Suneetha Kadiyala (Senior Lecturer, London School of Hygiene and Tropical Medicine), Carlo Cafiero (Senior Statistician, FAO), William Masters (Professor, Friedman School of Nutrition, Tufts University).

The assessment of global, regional and national efforts to reach the ambitious and important objectives of SDG2 will require innovations in approaches to define and measure hunger, food security and agricultural sustainability. Our discussions will focus on the following inter-related questions: For which dimensions do we already have indicators/indices/data; do we need 'new metrics' to measure all the dimensions of SDG2; what systems exist to measure these indicators, accurately and reliably; how complete and reliable is the global 'data architecture'; and where are the gaps?

Reflecting the expertise of the speakers chosen for this theme, our speakers will focus primarily on indicators relating to hunger, food security and nutrition. Agricultural sustainability will also receive some attention.

Our starting point for discussions relating to hunger and food security will be the strengths and limitations of the *Prevalence of Undernourishment* metric. As is well known, this relies on the compilation and analysis of food balance sheet data. While there have been significant improvements in these, gaps remain particularly with respect to the measurement of food storage, food losses and waste. (A limiting factor here is the level of agricultural statistical capacity; something that will be touched on in our discussions.) All speakers are unanimous in the view that the measurement of hunger and food security needs to go beyond for new metrics that go beyond kcal/person/day relative to requirements. They agree that much more attention should be given to measuring the quality of diets. However, they have interesting but divergent views as to how best this should be accomplished. Discussion will cover topics such as the merit of household and/or individual surveys of access and intake, experience-based food security scales (including FAO's FIES measure), the measurement of micronutrient intake, diet diversity and food consumption scores. It will also note the importance of distributional, seasonality and gender considerations. Innovative means of collecting such data– such as the use of market data and data collected via surveys using mobile phones will be discussed as will their integration with the global data architecture.

Improving nutrition is one of the SDG goals (SDG2) and is critical to achieving several other SDGs. On this topic, our session will begin with a brief summary of progress against the World

Health Assembly targets and then proceed to discuss the existing metrics and data sources to monitor nutrition indicators. Discussion points will include the following:

(1) For some indicators such as stunting and wasting, internationally agreed metrics exist. However, there are challenges associated with coverage and representativeness. How can the global data architecture be strengthened to address these?

(2) SDG2 refers to ending "all forms of malnutrition", thus giving (rightly) attention to micronutrient deficiencies. Coverage of these is patchy and often excludes vulnerable groups such as the elderly. Are there novel diagnostics that can be brought to bear here?

(3) Data gaps exist for socially excluded/vulnerable groups. How can these be addressed?

Our discussion of the measurement of agricultural sustainability will begin by noting that of the three SDG2 components, this is the most problematic. It will point to the importance of improving farm surveys as an important means of addressing this, alongside the need to define in a contextually specific manner what is meant by sustainable practices and who is considered a "small-scale" producer.

Underpinning SDG2 is a somewhat incomplete theory of change. SDG2 speaks to issues of increased in investment in agricultural R&D and improvements in the functioning of global food markets. But it gives primacy to enhanced productivity of small holder farmers and other often marginalized groups. While such emphasis is likely to contribute significantly to the achievement of SDG2 targets for hunger and food security in rural areas, it will not – by itself – address undernutrition which requires progress across a broad range of SDG2 targets including those relating to health, water and sanitation, and the status of women. It is also unclear whether it will be sufficient to meet the food needs of burgeoning urban populations.

The formulation of the SDGs was underpinned by much wider consultations than their predecessor, the MDGs. Reflecting on this, what scope is there for more participatory and qualitative methods in the measurement of SDG2.

The session

Our session will begin with presentations by Carlo Cafiero, Will Masters and Suneetha Kadiyala. Mr Cafiero and Mr Masters will focus largely on the measurement of hunger and food security (with some, but less, consideration of sustainable agriculture) while Ms Kadiyala will examine the measurement of nutrition. Each presentation will be followed by a short Q&A with the speaker on the technical points they raise. After the three presentations are complete, we will then have an open, structured discussion of these metrics, including their theory of change and the role of participatory processes.

———

Theme 2: Partnerships and development actors – Dealing with the increasing complexity of development processes

Chair: Simon Levine, Research Fellow, Humanitarian Policy Group, Overseas Development Institute

Speakers: Julia Betts (Independent consultant), Dorothy Lucks (Executive Director, SDF Global Pty Ltd), Cristina Tirado (Policy Adviser, UN Standing Committee on Nutrition, WG on Climate and Nutrition)

- For analytical (and evaluation) purposes it may be useful to further unpack the concept of partnership, for example according to stages of developing partnerships.
- With regard to the issue of complexity it would be useful to reflect some of the discourse how "complexity can be evaluated" that has been going on within the evaluation community for some time. It would be important to emphasize that the SDG agenda has been developed in recognition to the increasing "complexity" of sustainable development that involves a range of new partnerships. It is not just the complicated nature of those partnerships that poses a challenge for evaluation, but also the fact that this far-reaching transformation will depend on different and diverse partnerships that need to be evaluated.
- In order to cover the issue of complexity in a rounded manner, we recommend revisiting the guiding questions for this theme, in particular: How will addressing SDG 2 be a complex and multi-disciplinary challenge? What factors are involved and what causal pathways and assumptions? Based on that, which partners will need to be involved in order to address the multiple dimensions of the problem? What roles will be played by which actors at what levels from local to national to international across various sectors?

Apart from any technical differences in scope, targets or indicators between the SDGs and the MDGs, there is a clear political difference. SDGs are presented as being about local ownership and global, trying to blur a divide between rich countries and poor countries in the way they approach the poverty that exists in both. If the SDG framework is to be meaningful in setting a global political agenda on human development partners of the poorer countries, including the Rome-Based Agencies, will have to think differently about their roles in helping the countries of the world meet their SDG targets. In the past, despite a rhetoric of partnership, development agencies (and sometimes donors) have often seen their responsibilities as ensuring the successful delivery of specific projects or instruments with agreed outputs and objectives. The SDGs are predicated on the (belated) recognition that bringing about development has always been the outcome of the interrelationships of many actors, including different parts of governments, private companies, individuals and civil society. One of the overt aims of the

SDGs, it appears, is to transform this relationship into a more genuine partnership, in which the responsibility of any development partner is much broader, to support a country to achieving its SDGs.

The 'evaluability' of such a responsibility is only secondarily a matter for evaluators. A job description of 'supporting a country to meet its SDGs' is so broad and so undefined that it provides little help for any agency or individual in knowing what standards to set themselves and how to prioritize what to do. First and foremost, then, helping to understand the evaluability of the responsibility is to define the role to be undertaken, to help the agencies to know themselves what to do and how to hold themselves to account. The discussion on the challenges to evaluation of a role that is defined as a relationship between development partner and a country is a discussion about how best to undertake that role: how best to work with partner countries, while being accountable to the agency itself, to the partner country and also, critically, to the people whose lives need to improve if the SDGs are to be met.

Evaluators can help bring clarity on these questions for the agencies and for the individuals by addressing some of these challenges. Three main themes present themselves initially.

Success on SDGs at national and global level cannot be delivered by any one entity, neither a government nor a development agency. Progress always comes from the interaction of many actors, and the SDG initiative aims to bring as many actors as possible together into a partnership to maximize progress. The nature of 'partnership' though is not simple. It demands a significant overlap of objectives, and even the willingness of one organization to forego its own objectives for the sake of supporting another to meet theirs. Very different kinds of relationships will be needed between different actors, and these relationships will and should change over time. The consistent use of the word 'partnership' to describe the very limited and short-term collaboration that is needed for organizations to share funds jointly received does not help agencies to recognize the challenges inherent in supporting and managing a diverse array of such relationships. Those involved in the partnerships need standards by which to know whether or not they are on the right track in building those partnerships: their work and the partnerships cannot therefore be judged only by whether or not they deliver the SDGs by 2030. Building relationships takes years, and progress will look different – and perhaps should be assessed differently – for different kinds of partnership. Evaluators have a responsibility to help in clarifying whether the right partnerships are being developed and how the building of a partnership can be judged as a successful process.

Working with others in relationships almost inevitably brings in complexity. Everyone knows that no project proceeds as the simple unrolling of a predetermined plan and that the impacts of projects in real life do not follow a linear or predictable model of cause and effect. However, the bureaucracy of aid has usually encouraged everyone to continue working as if this were the case. Taking partnership seriously means ending that. 'Complexity' means that the challenges of meeting the SDGs are not only difficult (technically), but they involve dealing with unknown unknowns. It is impossible to know in advance what will need to be done by any one actor, because the outcomes of the interaction of so many actors cannot be predicted. Once politics (in its broadest sense) and power relations are acknowledged as a universal context for all human action, it must also be recognized that plans and responsibilities will be shaped over time. This presents huge challenges. Progress will be dependent upon establishing productive partnerships: but even the identity of the partnerships needed cannot be known in advance, since they will depend upon the development paths chosen by the interaction of different actors. Even when these are established, how can anyone or any agency work with a job description or terms of reference for an as yet undefined responsibility? Again, evaluators have to shine a light on this challenge by helping us all to see how it can be judged how well any agency or individual has managed the challenge of shaping their own – and others' – responsibilities, as well as in meeting them. Some progress is being made in learning "how to measure difficult to measure stuff", and complex processes and partnerships have begun to be evaluated with an eye to their complexity. Lessons gleaned from such progress should be a critical resource from the outset for those interested in participating in and helping shape these processes.

Partnerships and complexity lead to difficulties in establishing accountability. If my role is to contribute in partnership to a process for which I am not ultimately responsible, and if it is impossible to define in advance what specific tasks and outputs I should achieve, for what can I ultimately be held accountable? If it is an integral part of my role in a partnership to shape my own responsibilities as I go along, how can it be judged whether or not I have performed adequately or spent resources responsibly? The problem is sometimes posed as one of knowing my own role in bringing about change. Evaluators address part of the partnership dilemma through the language of contribution rather than attribution, though having a conceptual model for thinking about the challenge is not always the same as solving it. However the challenge is much deeper, and the help of evaluators is needed to see how we can all hold ourselves to account not just for the contribution we make to change, but also for the changes which in the course of our work we set out to bring.

These three challenges of partnership, complexity, accountability and made more difficult by the context for which the SGDs are designed. If the world were entirely populated by benevolent and collaborative institutions, there would have been no need to establish SDGs because food security would have been achieved long ago. Much of the problems of malnutrition, food insecurity and hunger are found in countries which are in protracted crises and conflict, or where there are weak governments, governments of contested legitimacy or predatory states. SDG2 cannot be met if we all ignore the fact that the greatest difficulty in meeting SDG2 will be precisely in the places where genuine partnerships are the least likely to occur, contexts are most uncertain, and standards of accountability are the most difficult to maintain. Business as usual on hunger, appropriate for many countries, will not work in the most difficult places. This poses challenges at national level: evaluators will have to evaluate in places where there is little reliable data, where it is probable that job descriptions ought to change far more regularly than they are changed in practice, where partners are often a party to the underlying causes of hunger. It also poses a global challenge. How far does it make sense to judge progress towards SDG2 as a global average? How far can lessons from one country be shared across all countries? How do we judge a global process if it is also argued that local specificity means that what happens in some countries has little to do with what happens in the rest of the world?

The session:

The three challenges of partnership, complexity and accountability will be dealt with respectively by Cristina Tirado, Julia Betts and Dorothy Lucks. They will treat the 'difficult places' challenge as a cross cutting issue.

A short plenary discussion will look at the big picture of all the challenges together.

The session will then break into small groups, each of which will be posed a more specific question. These have yet to be finalized.

Theme 3: National M&E systems and data availability – Building on the progress made and addressing existing (capacity) gaps

Chair: Carlos Barahona, Deputy Director, Statistical Services Centre, University of Reading

Speakers: Pietro Gennari (Director, Statistics Division, FAO), Ahmad Zaki Ansore (Director General, Implementation Coordination Unit of the Prime Minister's Department, Malaysia), Morten Jerven (Associate Professor, Simon Fraser University & Norwegian University of Life Sciences)

The process started with the agreement and announcement of the 17 Sustainable Development Goals and 169 targets the joint work of partners at national, international and sub-national level. This involves the negotiation of agendas, establishment of priorities, allocation of resources and learning from the results of activities that over the next 15 years have the potential to influence the path of development in the world.

There are a number of requirements and assumptions that are crucial in this process. An important assumption for our discussion has to do with the political willingness of national stakeholders to engage on the process of evaluating the SDGs, in particular SDG 2, and the ability of international institutions to work with, support and nudge the agendas of national governments.

If we were to assume that some level of agreement on agendas and priorities is reached that will allow monitoring and evaluating progress towards ending hunger, achieve food security and improved nutrition and promote sustainable agriculture, what are the challenges that are faced by the governments and other actors in the process of learning and adapting to improve results?

To monitor and evaluate SDG 2, consistent acquisition of data that will be converted into information is essential. This information needs to be available with the level of granularity that is useful to make decisions that allow for the diversity of 'relevant' contexts in each country. This requirement poses a challenge to the already established national systems in charge of gathering data and generating that information as they have been set up to fulfil information needs that were set well before the sustainable development goals were conceived. The definitions, nature of the data, the sampling frames, the frequency of collection and the geographies for which the estimates are prepared by national statistical systems are not 100% aligned with the requirements for SDG 2. The case of sustainable agriculture, i.e. the existence of sustainable food production systems that use resilient agricultural practices that increase productivity and production, that help maintain ecosystems and that strengthen capacity for adaptation to climate change, is particularly pertinent as currently there are very few routinely collected indicators about sustainable agriculture – an in most cases not even an agreed definition of it. The improvement of biodiversity, or its converse, the loss of biodiversity, is something that is not currently part of most national systems for M&E. The situation is less extreme in the case of nutrition, income and access to markets where routine data collection is already in place but the purpose for which the data are gathered, the way it is analysed and the processes the information feeds into tend to be different from the evaluation of SDG 2. For those indicators, the level of disaggregation that is currently possible for vulnerable groups also needs improvements.

Answering these challenges means establishing the capacity at institutional level to collect and analyze data that will serve as the basis for monitoring the direction of change of key agreed indicators and learn how different strategies help the interested parties to be more effective in the progress towards achieving the SDG 2. A tension that requires some attention is the lack of harmony in the granularity that different partners consider pertinent for the evaluation. National governments tend to be interested with whole countries, and that meets the interest of some international players. However, the activities and sphere of action of international agencies such as WFP, CG, FAO or IFAD are often targeted to sub-national geographies – and correctly so as the needs of specific groups of people often are priority within their agendas. At the same time the civil society, often represented by NGOs and sector-based organizations, tend to have interests in sub-sections of the population that may or may not be distributed throughout the national territories. Establishing the capacity to coordinate and integrate efforts by these actors for the collection of data, with the consequent advantages for analysis and lesson learning is not a trivial task, but may be at the center of a more effective approach that will help with understanding how we better target our efforts towards the achievement of SDG2.

Another challenge that should be borne in mind is the capacity, or interest, from many national governments for evaluation, and here lies the focus of our discussions. The need and ability to monitor exists in most national systems, but the requirement of evaluation implies a willingness to seek and weigh complex evidence; evidence that will lead to choosing among a set of options rather than unique answers, which in turn will be used to inform decisions to modify approaches and actions towards achieving the SDG2. In the world of national governments, evaluation is not as common as many would expect. Some of the opportunities that we may encounter include the extent to which international partners are able to identify and partner with parts of national M&E systems that are more susceptible and willing to build or enhance evaluation capacity. This has to do with identifying opportunities for building capacity within national systems and enabling the integration of capacities across players present in a national context.

Theme 4 - Demand for and Use of Evidence from Evaluation: Understanding the political economy of evidence and developing a joint evaluation agenda for SDG2.

Chair: Ian C. Davies, Credentialed Evaluator and a specialist in public governance, management, accountability and finance

Speakers: Thania de la Garza Navarrete (Deputy Director General of Evaluation, National Council for Evaluation of Social Development Policy [CONEVAL], Mexico), Osvaldo Feinstein (Independent consultant), Rossetti Nabbumba Nayenga (Deputy Head, Budget Monitoring and Accountability Unit, Ministry of Finance, Planning and Economic Development, Uganda)

The SDGs reflect a broad political consensus on a vision for our collective development and the goals attempt to capture our aspirations for a sustainable future. However they are not, as such, a technical blueprint for how to go about it and it would be misguided to treat the SDGs and their associated indicators as intended "impacts" at the end of a results chain. The SDGs are a political process.

Realizing the vision that the SDGs point towards will require innovation, experimentation, learning from failure as much as from success and acting on newly surfaced knowledge[1]. Simply put, we don't really know how to go about this collective endeavor the SDGs propose and in this simple fact lies a golden opportunity for evaluation to come of age as a politically engaged knowledge function.

Not that the SDGs change the fundamentals of sound evaluation and appropriate methodology. However, by their very nature and the principles that underpin them, they both challenge evaluation to evolve consistent with these and offer a global platform to rethink and reposition evaluation and its practice in development. The question then is not "how to evaluate the SDGs" but "what is evaluation's value proposition for a sustainable future?"

---

[1] See for example Harvard CID Working Paper No. 240 - Escaping Capability Traps through Problem Driven Iterative Adaptation (PDIA)Matt Andrews, Lant Pritchett, Michael Woolcock, June 2012.

And, from a political economy perspective, who assesses that value proposition? Who has legitimacy, rights and voice, in bringing their interests to bear on what evaluation addresses and how it is conducted? Who are the key stakeholders?

External financing partners, i.e. donors and lenders, that demand “evidence of results” for their accountability purposes? National governments of countries receiving external financing that require knowledge to inform policy processes? Civil society organizations that require truth and transparency about changes needed to achieve social justice? Citizens who have a right to know?

Questions such as these go to the heart of the political economy, or the politics, of “evidence”. That the word “evidence” itself, together with its conjoined terminological twin “results”, often dominate the discourse of donors and lenders around evaluation, is one of the clearest indications not only of the predominance of the accountability agenda but, most significantly, of whose interests and views are more important[2].

And it is through the use, some would argue the imposition, of terminologies that reflect particular world views and power relationships, that only some value constructs, and hence the evaluation approaches and methodologies deemed best suited to these, are acknowledged and the “evidence” that speaks to them considered “valid”.

Simply put, the terminology of “evidence” and “results” is used as an artefact[3] of power by those who have it in funder-recipient relationships. Being held to account for “delivering” against “targets” in a so-called logical framework in order to receive funds, e.g. “payment by results”, is a clear, and concrete expression of the artefact in use.

When “M&E” and evaluation “validate” these artefacts by founding their practices on them, they are in effect preserving and strengthening the established power differential. And doing so at times through a faddish and rote discourse peppered with old, e.g. “impact”, and refurbished, e.g. “resilience”, terms the manner of use of which belies a lack of intellectual rigor and scientific understanding.

This aspect of the political economy of evidence has largely characterized development and its evaluation however the SDGs offer the opportunity to set an evaluation agenda that addresses power relationships and changes fundamentally the role that evaluation plays within those power relationships. But to do so requires a deep reflection on, and an honest self-critique of,

[2] ‘Evidence’ and ‘results’ have a common intellectual heritage of ‘methodological individualism’ that economics shares with medicine; in this they differ from the holistic social sciences, which are concerned with relations between people and the culture and history that shape them. Rosalind Eyben in “The politics of evidence and results in international development” (2015).

[3] See Michel Foucault.

our evaluation approaches and practices[4], as well as a willingness of evaluation communities of practice to engage with political institutions .

A joint evaluation agenda consistent with the SDGs is first and foremost about evaluation and its communities of practice engaging in a process of reframing a global vision for evaluation as a politically engaged scientific endeavor to bring about social justice. And of developing explicit conceptual frameworks and their corresponding terminologies, in which to anchor the development and use of evaluation approaches and practices that engage on power relationships.

A joint evaluation agenda consistent with the SDGs should require of evaluation commissioners, such as the evaluation functions of the Rome based agencies that have brought us together, that they make explicit their vision of evaluation in the contexts within which they ply their craft and within which the evaluands evolve.

These visions should converge in a joint evaluation agenda that seeks to address rights rather than results and to communicate knowledge rather than evidence.

The realization of such an agenda also requires practical reconsiderations of accountability and learning frameworks, and their implications for the feasibility of evaluation approaches and practices that are relevant to the political opportunities and constraints that evaluation practitioners face in the field.

Accountability frameworks should rest on two fundamental principles:

- You are accountable only for things over which you have reasonable control.
- You are not expected to achieve the impossible.

They should make also explicit who is accountable to who, for what, how, when and where.

“Learning” frameworks should explain how insights acquired in the field through experimentation, trial and error, failures and successes in specific contexts, are fed back into the evaluative and evaluand systems, what is their intended use, e.g. single, double and triple loop learning, and what and how does evaluation contribute.

Central to this undertaking is a re-examination of the evaluation discourse in the context of the political economy and politics of development. The casting of the challenge to evaluation as one of supply and demand is in of itself revealing of a number of embedded paradigms, among

[4] “We can see our political institutions being robbed more and more of their democratic substance during the course of the technocratic adjustment to global market imperatives. Our capitalist democracies are about to shrink to mere façade democracies. These developments call for a scientifically informed enlightenment. But none of the pertinent scientific disciplines – neither economics nor political science or sociology – can, in and of themselves, provide this enlightenment. The diverse contributions of these disciplines have to be processed in the light of a critical self-understanding.” Jürgen Habermas (2015)

them a specific economic perspective, and of a tendency to simplify the complex inter-relationships and inter-plays of the multiple political forces that ultimately determine social justice and its sustainability.

Finally, it is in the koan of sustainability that the SDGs offer both a direct challenge to a static conception of "results", and by association "evidence", and a value based frame of reference for engaging in and on processes that are sustainable. For evaluation, focusing on processes will "create the space for appropriate methodological choices in support of processes of social transformation"[5].

The session:

The first hour or so will be devoted to presentations by the session panelists, with Rosetti Nabbumba from Uganda and Thania de la Garza Navarrete from Mexico discussing, inter alia, country perspectives, and Osvaldo Feinstein addressing, inter alia, demand for evaluation and its benefits for public accountability and learning. The format will allow for interactive discussion on the presentations between panelists and participants . The second hour or so will be organized around discussion groups with a view to surfacing key issues for evaluation from a political economy perspective and outlining considerations for how these may be addressed by evaluation. The last hour or so will seek to reap and record the outcomes of the discussions with a view to informing suggestions for follow up action.

[5] Rosalind Eyben in "The politics of evidence and results in international development" (2015)