Rob D. van den Berg
Cristina Magro
Silvia Salinas Mulder
Editors

EVALUATION FOR TRANSFORMATIONAL CHANGE

Opportunities and challenges for the Sustainable Development Goals

INTERNATIONAL DEVELOPMENT EVALUATION ASSOCIATION
IDEAS
KNOWLEDGE CAPACITY DEVELOPMENT NETWORKING

Evaluation for transformational change

Opportunities and challenges for the Sustainable Development Goals

Rob D. van den Berg
Cristina Magro
Silvia Salinas Mulder
Editors

INTERNATIONAL DEVELOPMENT EVALUATION ASSOCIATION
IDEAS
KNOWLEDGE CAPACITY DEVELOPMENT NETWORKING

EDITORS

Rob D. van den Berg
President, International Development Evaluation Association (IDEAS)
Visiting Professor, King's College London
Leidschendam, the Netherlands

Cristina Magro
Secretary-General, International Development Evaluation Association (IDEAS)
Independent consultant
Belo Horizonte, MG, Brazil

Silvia Salinas Mulder
Coordinator, Red de Seguimiento Evaluación y Sistematización de Latinoamérica y el Caribe (ReLAC)
Independent evaluator
La Paz, Bolivia

Exeter, United Kingdom
Email: ideascoordinator@gmail.com

Rob D. van den Berg, Cristina Magro, and Silvia Salinas Mulder (eds). 2019. *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals.* IDEAS, Exeter, UK.

ISBN (paper): 978-1-9999329-2-3
ISBN (electronic): 978-1-9999329-3-0
Kindle: 978-1-9999329-4-7

Design: Daniella Domingues

Evaluation for transformational change

Opportunities and challenges for the Sustainable Development Goals

Contents

Foreword

Gonzalo Hernández-Licona

Since 2015, the world has shown great enthusiasm for the Sustainable Development Agenda 2030. Everywhere we go, especially within governments, international agencies and academia, the topic is present in multiple colours. We need to applaud this enthusiasm. However, after more than four years since the launch of the Agenda, there are at least two threatening challenges for sustainable development.

The first one is that "the world is not on track for achieving most of the 169 targets that comprise the Goals", as it was underlined by the *Global Sustainable Development Report 2019*.[1]

While some goals are not on track, others are even going backwards, through rising inequalities, climate change, biodiversity loss, increasing waste from human activity, violent conflicts and related humanitarian crises causing the displacement of millions of people. We need to make more changes in the way we arrange our economies, our societies and our politics if we really want to have a better world in 2030. Business as usual will not do the trick. The most important argument of the report is that if we do not properly address the whole development system (or the multiple development systems underlined in the Agenda), if we do not take into account the various trade-offs and interlinkages between the Sustainable Development Goals (SDGs) and targets proposed in the Agenda, we will not be meeting most of the goals by 2030.

The second challenge for development is that the use of rigorous evidence (in a broad sense), is not the flavour of the month nowadays. Furthermore, the new fashion of spreading fake news by the media, politicians and one or two

1 Independent Group of Scientists appointed by the Secretary-General. *Global sustainable development report 2019. The future is now – science for achieving sustainable development.* New York: United Nations. https://tinyurl.com/y6y5qea7.

"professionals" confounds voters, news consumers and policy-makers from all latitudes, putting in danger crucial development outcomes.

Populist governments from the right, the left or even the centre, tend to put science and evidence aside, in favour of their own popular data. Their narratives are seductive and relate well to the situation of millions who have been left behind in terms of development over the past decades. Unfortunately, the public policy decisions of autocrats who believe they know best, will not be the solution for development either. Quite the contrary.

Is there anything we can do to tackle these two challenges? I guess we all have many things to accomplish on various fronts, but there is one actor that has big possibilities to help sustainable development: Evaluation.

For decades, there have been three triggers of evaluation around the globe: democracies, the need for efficiency and technical advances. Not only in developed countries evaluation processes increased substantially, in many developing countries – India, Mexico, Colombia, Chile, the Philippines, China, South Africa, Uganda – evaluations are now part of public policy.

Two or three decades ago we only witnessed very few evaluations in each of these countries, basically sponsored by international agencies. Today we can see monitoring and evaluation (M&E) systems in several countries, where many evaluations – impact, process, design, consistency – are developed every year. Many of these evaluations have helped several governments, national and local, to improve their projects and programs.

But if we would like evaluations to fully address the present development challenges, we need to go beyond what we already achieved. Business as usual in evaluations is not an option, either. This book is an excellent guide for these new paths for evaluation in the midst of the 2030 Agenda.

If we would like to avoid the silos-type public policy, we need evaluations that go beyond single programmes and projects. We need dynamic evaluations able to evaluate strategies and policies. We need the strength and innovation of young evaluators. More importantly, evaluators are urged to evaluate transformational change.

Climate change and environmental problems cannot be solved without thinking about systems and without addressing explicit development trade-offs. These problems can be understood better if we think about transformational change

("deep, systemic, and sustainable changes with large-scale impacts in a significant area of concern", as defined in chapter 7 of this book). Good impact evaluations in the past showed, for example, how a single budgetary programme had good results on education attainment in rural areas in Mexico. But these rigorous impact evaluations cannot address complex systems like the ones needed to understand what we can do, and need to do very soon, in order to improve the quality of our environment. Systems evaluations for transformational change are the answer, according to this important book.

If we apply good evaluations to the development process of the 2030 Agenda, we may have more chances of meeting the goals we agreed more than four years ago. If through democracy we can push governments to implement good evaluations as a common practice, then we will be able to improve their annual Voluntary National Reports (VNRs) and thus their SDG implementation. We need to accept that implementing the 2030 Agenda is not easy and most countries are not fully committed to doing so, which can be seen reading their VNRs. They are good looking reports to impress the UN and other countries. But as we have seen, these reports don't match with the reality of their development situation.

New and better evaluations can help us with fake news? Yes, up to a point. Only if we combine them with more democracy and transparency, that is, with the help of civil society defending important institutions. Chapter 6 of the book identifies the need to use social values for transformative evaluations. Evaluators should find what type of values (cultural values for example) are important in societies, especially in countries facing severe problems, in order to make evaluations more relevant and reliable for all stakeholders.

I propose to apply the same thinking to sustain and strengthen evaluation systems around the world. Whenever we find politicians, in Northern or Southern countries, neglecting rigorous information coming from solid statistical or evaluating institutions, civil society should protect these institutions as a core element of the countries' democracy. Populist governments do not have the right, like the Taliban, to destroy the countries' heritage built for centuries. The advances shown on evaluation over the past decades, together with better evaluation methods and narratives to address the 2030 Agenda, should stand strong despite any challenge. The wellbeing of millions deserves this.

Acknowledgements

The editors would like to thank all authors for their contributions, their willingness to deliver under strenuous deadlines, and to finalize their chapters under duress! But most of all we thank them for their experiences, thoughts, ideas and the challenges and opportunities they formulated for IDEAS and for the global evaluation community. All authors have been identified by their full names without titles, thus following the United Nations tradition of an equal perspective on all contributors. The bios of the authors contain further information on their background and achievements, in the session Contributors of the book.

A special thanks to Gonzalo Hernández-Licona, one of our keynote speakers at the 2017 Global Assembly, who just went through a difficult time when a new President took office in Mexico leading to a separation of Gonzalo from the evaluation institute CONEVAL, to which he dedicated his professional services and passion for many years. Nevertheless, Gonzalo generously accepted to deliver a foreword to this book.

This publication has been made possible through a grant from Universalia in Montreal, Canada (www.universalia.com). The book, including cover and lay-out, both in its printed, PDF, EPub and Kindle versions, has been designed and formatted by Daniella Domingues, Brazil, to whom we express our deepest gratitude, for having followed us in the adventure of bringing this book to the public in such a tight schedule.

References are currently a challenge, as the internet is slowly displacing the library as an archive for documents, reports, journals and books. In the course of working on this book IDEAS has changed its reference guidance to accommodate this development. Internet addresses have been added where possible while they used to be accompanied by formulations like "available at" and rounded off with "accessed at this date", we have decided to abandon these formulations, as they have become generally understood. If there is a web address, the publication can be found there or was there when last checked in the preparation of the publication, which in the case of this book was August-September 2019.

In the strenuous work of reviewing the references of each chapter of the book, and of building a reference style guide for IDEAS that required several changes while working on draft chapters, we express special thanks to René de Winter.

Finally, we thank René de Winter, Jasmijn van den Berg and Daniel Svoboda for the support in the launch of the book at the Conference "Evaluation for Transformative Change: Bringing experiences of the Global South to the Global North", which includes the 2019 IDEAS Global Assembly and the Third International Conference on Evaluating Environment and Development, in October 2019, in Prague, Czech Republic.

Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder

—

Addendum to the version for Printing on demand, Kindle and ePub available as from November 2019.

This book contains contemporary issues to be discussed at the Global Assembly of IDEAS in Prague, 2-4 October 2019. This event has taken place and the challenges of evaluating transformational change have been discussed. We hope that the essays in this book will continue to inform discussions at future events of the global evaluation community and in meetings of communities of practice and discussion forums, wherever they would enrich thinking and offer a good starting point for further deliberation.

The editors, November 2019.

Introduction: Bringing challenges for evaluation for transformational change from Guanajuato to Prague

Rob D. van den Berg, Cristina Magro, and Silvia Salinas Mulder

ABSTRACT. In this introductory chapter the editors set the stage for a discussion about how evaluation could contribute to transformational change. A historical overview of the international support for and understanding of Agenda 2030 is presented, and the book is placed in the history of previous publications edited by IDEAS, showing the continuous commitment of the Association to discussing subjects which are up to date in the world of development evaluation. Each chapter is individually presented and articulated with the threads that compose this interlinked history.

Aspirations for transformational change

Resolutions of the United Nations tend to be dry, bureaucratic and often difficult to read. Boring stuff, in other words. On 25 September 2015 the United Nations members adopted unanimously a document (A/Res/70/1) that promised dramatic change: "*Transforming* our world: the 2030 Agenda for Development" (emphasis ours). This high ambition is followed in the preamble of the resolution in five promises that read more as rallying cries than budgetary and regulatory deliberations. We quote them here as they are astounding in their vision, in their clarity and in describing which transformation is required, with broad strokes rather than in fine detail:

> ***People*** We are determined to end poverty and hunger, in all their forms and dimensions, and to ensure that all human beings can fulfil their potential in dignity and equality and in a healthy environment.
>
> ***Planet*** We are determined to protect the planet from degradation, including through sustainable consumption and production, sustainably managing its natural resources and taking urgent action on climate change, so that it can support the needs of the present and future generations.
>
> ***Prosperity*** We are determined to ensure that all human beings can enjoy prosperous and fulfilling lives and that economic, social and technological progress occurs in harmony with nature.
>
> ***Peace*** We are determined to foster peaceful, just and inclusive societies which are free from fear and violence. There can be no sustainable development without peace and no peace without sustainable development.
>
> ***Partnership*** We are determined to mobilize the means required to implement this Agenda through a revitalised Global Partnership for Sustainable Development, based on a spirit of strengthened global solidarity, focussed in particular on the needs of the poorest and most vulnerable and with the participation of all countries, all stakeholders and all people.
>
> (United Nations 2015, 2)

It does not stop there. Rather than the time-honoured silo approach of development, where economists would rule the economy, health professionals would rule health care, and education professionals education, engineers infrastructure and evaluators evaluation (to name just a few), Agenda 2030 calls for integrative action:

The interlinkages and integrated nature of the Sustainable Development Goals are of crucial importance in ensuring that the purpose of the new Agenda is realised. If we realize our ambitions across the full extent of the Agenda, the lives of all will be profoundly improved and our world will be transformed for the better (United Nations 2015, 2).

The gradually accepted urgency of the challenges posed by Agenda 2030

The ambition of Agenda 2030 and the embedded Sustainable Development Goals (SDGs) met at first with ridicule and sarcasm, perhaps best expressed by William (Bill) Easterly in an opinion piece in Foreign Policy. In his opinion "only the U.N. could have come up with a document so worthless" (Easterly 2015). Ridicule often overshoots its mark and Bill managed to do this in a spectacular way. The Millennium Development Goals, hailed by Bill as "precise and measurable", were actually written and prepared by UN officials. The Sustainable Development Goals were written and prepared by development experts from a great variety of organisations and countries in a carefully prepared and wide-reaching participatory process with representatives of many organisations, universities, governments, groups, civil society organisations, and the private sector. The process itself was highly applauded and led to a document in which no letter or comma was changed by the UN. For Bill this would not be a problem, as on top of the UN he has also often ridiculed experts. The designation of the SDGs as worthless actually came from the Economist, who earlier on in the process described the emerging document that would become Agenda 2030 as "The 169 commandments: the proposed sustainable development goals would be worse than useless" (Economist 2015).

On the other hand, the SDGs were defended by writers like Homi Kharas in a blog on the Brookings website, where he revealed to expect "great outcomes" from the SDGs, as they would set the tone and substance of the debate in the years to come and thus focus the world on searching for the solutions to the problems noted in Agenda 2030 (Kharas 2015). In general, the SDGs were identified as "aspirational" rather than the concrete and measurable approach in the Millennium Development Goals. The SDGs could be described as the Universal Declaration of Human Rights for Development purposes. The Human Rights declaration is clearly aspirational, is highly difficult to "measure" and often lacks precision. It has been accused of bias and being incomplete, yet not many would want to see it abolished, as its aspirations are considered truly worthwhile, even if its implementation in the world

is incomplete and sometimes controversial. Instead, efforts have often focused on expanding the Universal Declaration, to include indigenous perspectives, gender identities, and development from the perspective of the Global South. While many if not all accept the Human Rights declaration as "worthy" of setting goals for our societies to achieve, it could be argued that Agenda 2030 addresses some of the criticism levelled at the Universal Declaration and aims to add to it rather than to subtract. It is not difficult to see that the SDGs set similar "worthy" goals for global development as the Universal Declaration sets for human rights.

Opinions like the ones of the Economist and William Easterly are mirrored and magnified in and by the emerging populist movements since 2015, who in general tend to disregard and disqualify Agenda 2030 and the Sustainable Development Goals and see them as a global conspiracy against the "people" and the sovereignty of countries, ethnicities, cultures and religions. The Economist has moved in the other direction and is increasingly willing to recognize the dangers of populist movements for the future of humankind on our planet. While they may not have become avid fans of the SDGs, their articles increasingly focus on the need for transformational change to prevent the collapse of our societies, economies and our habitat. In its 175th anniversary edition, the Economist raised the need for liberalism, identified as "a universal commitment to individual dignity, open markets, limited government and a faith in human progress brought about by debate and reform", to wake up out of its complacency and become radical once again. Their verdict was eerily similar to the condemnations of the populists: "the ruling class lives in a bubble". Most importantly, they asked liberals to side with the people "against the patricians" to tackle the growing problems (Manifesto 2018).

True to its words, the Economist followed up with leaders about "Crude awakening. The truth about Big Oil and climate change", "The rise of millennial socialism" and "The global crisis in conservatism", "Deathwatch for the Amazon" and "Democracy's enemy within".[1] While the Economist is just one opinion publication and not the measure of change in world opinion, it nevertheless demonstrates a shift from a publication supporting the neoliberal, Washington-consensus style "business as usual" towards further globalisation to a renewed liberal vision that has taken over and adopted many of the concerns as expressed in Agenda 2030. The Sustainable Development Goals were developed partly in protest against the dominant neo-liberal perspectives of the Millennium Development Goals — it is

1 Respectively the editions of 9 February 2019, 16 February 2019, 6 July 2019, 3 August 2019, 31 August 2019.

good to see that one of the global guardians of liberalism came around to support at least part of Agenda 2030.

Transformational change in development

In development two kinds of transformational change were recognized. Societal change had to be transformational to ensure a move towards gender equality and to equity in societies, to enable them to become inclusive. This was strongly supported from feminist perspectives and emancipation movements, from indigenous peoples to cultural and religious minorities. Gender equality was thus the sparkle for a much broader and stronger view of how radically segregationist many societies have historically been and for focused action against discrimination. Moreover, it also opened the world's eyes to a broad array of mistreatments of women and children throughout various cultures, from mutilations to culturally accepted rape.

A second perspective on transformational change emerged from the development and environment arena. From the Earth summit in 1992 in Rio de Janeiro onwards the sustainability of economies and societies that degraded their environment was put into question and the need for a transformation of interactions with ecosystems and the environment was clearly seen and promoted, often without much success. When it became clearer that humanity had reached planetary boundaries in its use of resources, the urge for transformational change became most visible in the fight against climate change with the call for a radical transformation of the energy sector from non-renewable to renewable energies.

The third drive towards transformational change came from management theory, operations research and the business advisory consultancies. McKinsey describes it as a "key source of competitive advantage" and aims to deliver it to their business clients.[2] The Business Directory defines it as:

> A shift in the business culture of an organization resulting from a change in the underlying strategy and processes that the organization has used in the past. A transformational change is designed to be organization-wide and is enacted over a period of time.[3]

2 See their website at https://tinyurl.com/y2eba8db.
3 See http://www.businessdictionary.com/definition/transformational-change.html.

Key elements are the integral, systemic character of the process – it is not just changing a product, or tackling the financial situation of the company, but a general overhaul that will enable the business to compete in markets, rethink the company's mission and vision to ultimately transform the markets in which they operate. The shining example of how this should be done, and how it has been done to an amazing extent, is the "digital transformation" of the world through the efforts of the Silicon Valley top companies, like Apple, Facebook, Twitter, Netflix, Paypal and others.[4] Their transformational successes have been matched by companies like Google, Microsoft and Amazon, and it is this transformational success that became a third factor in introducing it in Agenda 2030, as Agenda 2030 included the private sector as an important partner in development.

These three transformational perspectives were easy to transpose to other arenas of Agenda 2030. Peace requires a complete transformation from situations of conflict and violence to sustainable peaceful relations between warring factions and societies. Solving poverty, health care, education and many other areas identified in Agenda 2030 would also require transformational change. Furthermore, it also emerged that these transformations would need to take place throughout the world: the distinction between "developing" and "developed" countries was redefined: all need transformational changes to survive and to become truly sustainable and inclusive, "leaving no-one behind".

Consequences for the global evaluation community

The consequences of all these profound changes in the development paradigm were not immediately clear, especially in the global evaluation community, which to a large extent continued "business as usual", playing its role in monitoring, evaluation and learning in countries, in international and national development organisations and to a lesser extent in the private sector and in civil society organisations. In 2015, shortly after the Sustainable Development Goals were adopted by the General Assembly of the UN, IDEAS held its Global Assembly in Bangkok on Evaluation for Sustainable Development, to explore experiences and approaches regarding the evaluation of the underlying concept of sustainability in Agenda 2030. This conference took place in parallel with the 4th National Evaluation Capacities conference of UNDP, which focused on Evaluation for

4 See https://siliconvalley.center/blog/digital-transformation/.

the SDGs. While representatives of the global evaluation community discussed sustainable development in the IDEAS Global Assembly, the representatives of governments and evaluation institutions discussed how countries would approach the role of evaluation in reporting on the Sustainable Development Goals. This was very appropriate, as for the first time in the history of the UN evaluation is called upon to play a role in reporting on achievements of countries. Paragraph 74 of the UN resolution on Agenda 2030 states that country-led evaluation and evaluation programmes will provide evaluative evidence and inform follow-up processes at all levels (UN 2015, 32). This became a key focus of the 2017 IDEAS Global Assembly on evaluation for the Sustainable Development Goals, with added emphasis on the Latin American experiences through the conferences of ReLAC and RedLACME, together forming the Joint Conference of Guanajuato, Mexico.

Just before the Guanajuato conference, the IDEAS book *Evaluation for Agenda 2030. Providing evidence on progress and sustainability* was published and made available through the website of IDEAS. This publication contais chapters based on sessions in the Bangkok conference and was supported and made possible by UNDP. The administrator of UNDP, Achim Steiner, wrote the foreword. This book followed a tradition started by Ray C. Rist as President of IDEAS in the period 2008-2014. Three books were published by Ray and his co-editors (Marie-Helene Boily and Frederic Martin) on Global Assemblies through the World Bank press, all dealing with key topics that would be included in the SDGs: the role of evaluation in strengthening governance (Rist 2011), dealing with crises that endanger our future (Rist 2013) and tackling poverty and inequality (Rist 2013). There is thus a natural progression to the SDGs and the Agenda 2030 publication. Many of the subjects that started to play a key role in the discussion about evaluation for the SDGs were already identified and discussed in the previous three IDEAS books.

After the Guanajuato conference a book was planned along the lines of the *Agenda 2030* publication. It gradually became apparent that the call for transformations formulated in the title of *Agenda 2030* was becoming more urgent and more insistent. Donors started to ask international organisations to provide proof of transformational impact, and countries began to discuss transformational changes that would be needed to make progress towards the SDGs. The global "buzz" on transformational change increased dramatically. As Magro and Van den Berg argue in chapter 8, the perceived need for transformational change is shared throughout the political spectrum, from left to right, from traditional to populist parties. Disagreements would be mostly on who should pay for these changes, and whether crises could be averted by them or whether the transformation would need to ensure societies would adapt to the new situation.

When IDEAS decided to focus the 2019 Global Assembly on evaluation for transformational change, it also became clear that a book on evaluation for Agenda 2030 would itself need to be transformed to a book on transformational change. The motivation was shared with Silvia Salinas Mulder, who was elected in Guanajuato as ReLAC Coordinator, and who expressed that the focus of ReLAC for the 2018-2019 period was going to be on leaving no one behind, boosting evaluations' transformative potential in the region and strengthening the South-North dialogue towards the SDGs. And the history of this co-edited book began.

Presenting "learned arguments" for what needs to be done

Together we decided that what we wanted to deliver, on the basis of the discussions in Guanajuato, and the increasing attention for transformational change, would be essays that would pose the opportunities and challenges for evaluation to contribute to transformational change. We approached the authors of this volume with the promise that they could write an essay rather than an academic article that would be based on research or evaluative evidence. An essay is not necessarily a lesser effort than an academic chapter. It is used often as a way to allow the development of an argument regarding a topic worth discussing. Essays are seen as striving towards a "learned argument" that exhibits insight, mastery, and promise of future research. We invited our authors to develop a perspective on the opportunities and challenges of evaluation for the SDGs as seen from a transformational change perspective. This challenge has been met by them in surprising ways.

Dynamic evaluation

Osvaldo Feinstein in the second chapter of this book introduces the concept of Dynamic Evaluation. He argues that evaluation needs to adopt a new paradigm of undertaking evaluations, if they are to provide evidence on and influencing of transformational change. As we noted above, many of the elements of this emerging new paradigm have been visible in various efforts already for a number of years. His contribution is to argue for bringing these elements together to enable the global evaluation community to further develop this model and make it truly responding to the needs for transformational change. He systematically goes through key aspects of the proposed dynamic evaluation paradigm, starting off with an evaluation agenda that is relevant and significant for the transformational change required to bring sustainability to the world. He argues for the need to incorporate multiple methods and techniques into dynamic evaluations as a matter

of principle, given the interlinked nature of societal, economic and environmental problems. The promising nature in this regard of ICT and "big data" is raised. Scaling-up in his view needs to become an integral part of what the evaluation would look for, and if it does not happen, what the barriers are against it. He champions quasi-real-time evaluation and argues that dynamic evaluations need to be politically sensitive, to ensure the evidence brought to the political domain will strengthen democratic processes of introducing transformational change.

Complexity and context are of key importance in Osvaldo's perspective, as he sees the SDGs as components of a system; he argues for the use of contextual tools like the Cynefin framework. Given these complexities, he argues that dynamic evaluation should move towards triple loop learning efforts, which he poses as "transformative learning", leading to a change in our understanding of the problems that the SDGs aim to tackle. For this it is necessary to include evaluation of the policy dialogue and of the societal issues that require transformational change. His chapter ends with the challenges for our profession: how can we build the capacities to undertake dynamic evaluations? He points to the importance of including science in evaluation and supports the efforts to introduce an International Evaluation Academy, as will be discussed at the 2019 IDEAS Global Assembly in Prague (2-4 October 2019).

The Youth revolution

One essential ingredient in tackling the future perspective in evaluations focusing on transformational change is to include youth. Bianca Montrosse-Moorhead, Khalil Bitral, Josette Arévalo and Antonina Rishko-Porcescu argue in their provocative and challenging essay in chapter 3 that the revolution towards dynamic evaluations focusing on transformational change by necessity means a revolution towards participatory youth evaluation as well. They establish once again that youth's stake in the future is higher than that of older people, and claim their place in evaluations relevant and significant for that future. They argue for a move from evaluation that observes youth to evaluation with and by youth and see promise in the appointment of a UN Youth Envoy and the adoption of a Youth Strategy by the UN.

While they recognize that youth in general represents all ends of the political and social spectrums, from laissez-fair to radicalism, they point to the values that youth would bring to evaluation for transformational change, especially where youth is active in grassroots movements for enhanced social justice, equality, equity, environmental sustainability, peace and so on. Moreover, many of these

movements employ new methods of organisation and of engagement, that could considerably enhance evaluative efforts. They point to the skills that youth would bring to evaluation, with enhanced use of the new media, eLearning, and their great skills in communicating through these new media. Lastly, there is a growing supply of Young and Emerging Evaluators (YEE) to contribute to evaluations for transformational change. They end by posing the challenge to the global evaluation community: where will you stand on this emerging youth revolution? Will history put you on the right side?

Transforming national monitoring and evaluation (M&E) systems in Africa and Latin America

For transformational change at the national level, in line with national priorities derived from the SDGs, the emerging M&E systems in many countries face great challenges. Some emerging experiences and lessons are unveiled by Abdoulaye Gounou for Africa and Gabriela Pérez Yarahuán for Latin America in chapter 4. While the section on Africa reveals the great variety in stages of development of a national M&E system, the Twende Mbele cooperation between Benin, South Africa and Uganda brings hope to the continent and demonstrates how a system can be developed that would be indigenous to Africa and relevant and significant for the development priorities of these countries. At the same time, it becomes clear that building capacity, institutionalising a system and developing the interaction between evaluative evidence and political and development decisions in line with the national SDGs' priorities is a major challenge in itself, not easily followed by integration of perspectives on transformational change. However, a key element will be how Voluntary National Reviews of progress towards the Sustainable Development Goals will incorporate evaluative evidence, including on transformational change.

In Latin America changes have taken place in many countries and most governments now recognize the need for evaluative evidence and learning in national systems, as requirements in the policy-budget cycle, as argued by Gabriela Pérez. An assessment of CLEAR-LAC of the integration of evaluation in national systems shows great varieties of progress throughout the region. Yet at the same time a recent study also reveals a steady growth of governmental evaluation efforts with an almost exponential curve (see chapter 4). Pérez and Gounou see evidence of innovation and piloting that is leading to systems that are not just a mimicry of well-established systems in the Global North, but systems that reflect local realities and challenges and include adaptive management that turns these systems into experiences of the

Global South.

Transforming M&E in the Caribbean and the Pacific

A formidable team of writers explores the transformative agenda for evaluation in Small Island Development States (SIDS) with emphasis on the Caribbean and the Pacific in our chapter 5. SIDS are at the frontline of impact of many of humanity's global crises. Climate change affects small islands in dramatic ways, sometimes leading to forced resettlement and loss of livelihood perspectives. The vulnerability of these islands on issues of food security, housing, health, education and security poses additional barriers for progress towards the SDGs. For evaluation to play a role in this regard, the islands are confronted with highly challenging capacity issues – leading to problems in how evaluation can take shape. Through sharing the lessons learned and experiences the SIDS aim to build an evaluation culture, get support at the political level, to enable further work on the role of evaluators and the systems in which evaluation can play a role.

While there are differences between the regions, which becomes clear in the sections led by Viliamu Iese on the Pacific and led by Lennise Baptiste on the Caribbean, there are also many similarities, and this has enabled the team to draw conclusions for SIDS in general, pointing to a lack of evaluation culture in government and in local communities, where too often evaluation is seen as a donor's pre-occupation and there is a negative perception of what evaluation can contribute. There is a lack of capacity in governments to deal with evaluation, which means most evaluators work for donors rather than their own island states. Both within and between the regions there is a history of collaboration which may provide solutions. The writers conclude that leadership is needed, that could support evaluation and its role, with a strong link to country-owned data collection, and sufficient attention for competencies and professionalisation. They propose to work on a repository of evaluation reports and sharing of lessons learned as important steps towards internalising evaluation as support for the transformative challenges that the two regions have to face.

Value-based evaluations in contexts of fragility, conflict and violence

The essay of Inga-Lill Aronsson and Hur Hassnain on value-based evaluations as our chapter 6 focuses on a key issue in the transformation from situations of conflict and violence towards sustainable peace: the issue of history, of deeply rooted values in communities and institutions, and of the role of the manifestation of heritage that

enables a discussion of these values, aiming at a permanent solution of conflicts. They explore the role of evaluation through a case study of Sierra Leone and its heritage, with a troubled past rooted in the slave trade, colonialism, independence, civil war and other crises such as the outbreak of Ebola. They conclude that efforts to tackle outstanding issues need to take heritage into account to solve them.

In order for evaluation to support such a process, they propose value-based evaluation, which would put vulnerability and respect at the core of the evaluation. The narratives of the past should play a crucial role in how society sees its transformation. Heritage has to be studied in all its manifestations (physical, institutional and informal) to reach understanding of the situation and enable a conversion of this into indicators and targets for a systemic evaluation. The challenge is for evaluations to bring out local and national values driven from heritage perspectives and to connect these to a sustainable peaceful future.

Evaluations of transformational environmental funds

Three major global environmental funds, two of which are focused on climate change (the Climate Investment Funds and the Green Climate Fund) and the third of which encompasses all environment and development linkages (the Global Environment Facility) are engaged with evaluation of transformational change. Their experiences are revealed in our chapter 7. The section on the Green Climate Fund (GCF), written by Jyotsna Puri and Archi Rastogi, explores paradigm shifts to come to a better understanding of how these could be evaluated to support the investment plans supported by the GCF, building on a framework developed by the Independent Evaluation Group of the World Bank. This framework, focusing on relevance, depth of change, scale of change and sustainability will guide future efforts of the Independent Evaluation Unit of the GCF.

The Independent Evaluation Office of the Global Environment Facility (GEF) developed a theory of transformational change for GEF supported interventions and evaluated a sample of GEF projects to test this out. In the section written by Juha I. Uitto, Geeta Batra and Kseniya Temnenko clear indications of transformational change were found. Key factors turned out to be a transformative ambition in the design, a focus on market and system changes through policies, the inclusion of mechanisms for financial sustainability and a high quality of implementation. A hopeful note is that although bigger projects scored well, even smaller projects with limited duration were able to support transformational change through tackling key barriers and involving key stakeholders.

The Climate Investment Funds (CIF) report on an initiative of their Evaluation & Learning programme: the Transformational Change Learning Partnership (TCLP). In this section, written by Anna Williams and Joe Dickman, a framework for evaluation of transformational change was used that is similar to the one proposed by the GCF, with one change: "depth of change" (GCF) was translated into "systemic change", which is more in line with the GEF. The considerable work that was done by and for the TCLP led to a first attempt at measuring signals of the advancement of change. Following the change brought about by CIF supported investment programmes in countries led to the conclusion that in all of these programmes advancement could be demonstrated, with especially clean technology investments advancing toward transformational change, whereas other programmes were still in intermediary stages.

Systems evaluations for transformational change

In the last chapter by Cristina Magro and Rob D. van den Berg, systems thinking is introduced as a prerequisite to understanding how evaluations can include systems analysis and use the appropriate approaches and methods to focus on transformational change of systems. They develop the argument that the global crises that the SDGs aim to address are systemic in nature and need to be evaluated and solved as systems issues. They explore systems thinking to bring clarity to debates about climate change, use of natural resources in societies and our economies, gender and equity issues as well as issues of war and peace.

Systems analysis is a radically different perspective from the reductionist approach and linear reasoning in which micro-actions cause changes at macro levels. Experiments with "what works" at a local level is not "what works" at the macro level, where higher level causal mechanisms take over. They point to the need to include Bayesian statistical approaches to include forecasting of developments (if transformational change is the purpose, we need to be able to forecast what the best route is towards this change) and to include new mathematical analytics, such as power laws and chaos mathematics, in our toolbox as evaluators. They end their chapter with a call to support the proposed Independent Evaluation Academy as they echo Osvaldo Feinstein's call for strong involvement of science and scientific tools to support evaluation for transformational change.

Final remarks

This book is by no means comprehensive or exhaustive. While societies will need to transform themselves to bring, for example, gender and ethnic perspectives to bear on societal and economic structures, and fundamentally different relations between humanity and the environment, the discussion of how evaluation can support these processes is under way in various meetings, conferences and books produced in the field. The same can be said about transformations in other areas targeted by the SDGs, of which the chapters of this book bring us a promising flavour.

Most important of all is that the chapters gathered in this book bring "learned arguments" for transformational change of our evaluations for transformational change… They explore how this would work out for supporting the SDGs, from areas of work that are advanced in this regard (chapter 7) to areas where progress still encounters barriers (chapters 4 and 5) or even major challenges (chapter 6). They propose a new approach (chapter 2), a new thinking and a new toolbox (chapter 8), while calling for a revolutionary transformation of the evaluation profession itself (chapter 3). We believe that this collection of essays will contribute to a further discussion of how evaluation can support transformational change in a world that needs this more than ever.

References

The 169 commandments. The proposed sustainable development goals would be worse than useless. *Economist*. [print version, leader] 26 March 2015 https://tinyurl.com/y22gr4a8.

Easterly, William 2015. The SDGs should stand for Senseless, Dreamy, Garbled. *Foreign Policy* September 28, 2015. Available at https://tinyurl.com/y2lmtp2x.

Kharas, Homi 2015. I expect great outcomes from the Sustainable Development Goals. *Brookings* blog, October 8, 2015. https://tinyurl.com/y2eywyyz.

A Manifesto. The Economist at 175. *Economist*. 15 September 2018. 175th anniversary edition.

Rist, Ray C., Marie-Helene Boily & Frederic Martin (eds). 2011. *Influencing change. Building evaluation capacity to strengthen governance.* Washington DC: The World Bank.

Rist, Ray C., Marie-Helene Boily & Frederic Martin (eds). 2013. *Development evaluation in times of turbulence. Dealing with crises that endanger our future.* Washington DC: The World Bank.

Rist, Ray C., Marie-Helene Boily & Frederic Martin (eds). 2015. *Poverty, Inequality, and Evaluation. Changing perspectives.* Washington DC: The World Bank.

United Nations 2015. Resolution adopted by the General Assembly on 25 September 2015. New York: UN A/RES/70/1. 21 October 2015.

Van den Berg, Rob D., Indran Naidoo, and Susan D. Tamondong (eds) 2017. *Evaluation for Agenda 2030. Providing evidence on progress and sustainability.* Exeter: IDEAS https://ideas-global.org/evaluation-for-agenda-2030/.

CHAPTER 2

Dynamic evaluation for transformational change

Osvaldo Feinstein[1]

ABSTRACT. This chapter presents a type of evaluation which tries to contribute to transformational change, i.e., dynamic evaluations. The focus is on transformational change of society, recognizing that transformational change can happen in many contexts, but that transformations for achieving the Sustainable Development Goals should be changes of society. Dynamic evaluations would aim to evaluate efforts to change societies. Key issues considered are the agenda for these evaluations, methods and techniques, scaling-up, quasi real-time evaluations and third-loop learning. Implications for dynamic evaluation capacity development are also discussed.

1 I am grateful to Rob D. van den Berg who encouraged me to write this chapter and provided excellent comments and suggestions; to Robert Picciotto for his valuable comments on the first version of this essay; and to Silvia Salinas Mulder and Cristina Magro for their comments.

Introduction

This chapter attempts to introduce, discuss and characterize *dynamic evaluations*, i.e. evaluations that aim to contribute to transformational change. This is made by identifying key issues that need to be tackled and providing guidance to deal with them.

If evaluations are expected to contribute to transformational change of societies, as it is the case in the context of the Sustainable Development Goals (SDGs), some challenges are to be faced. For example, discussions on the relative weight of learning vis-à-vis accountability still happen among evaluators, who assume that there is a trade-off between these roles. At the same time, others argue for an alternative view of complementarity or synergy between those two functions. The usual ways of formulating the roles of evaluation in terms of learning and accountability are no longer adequate. Evaluation and evaluators need to go beyond this traditional formulation and discussion if they are to effectively contribute to transformational change. This essay highlights some of those challenges and suggests ways to address them.

A first challenge is the definition of *transformational change of society* and how to incorporate it into the evaluation practice. A second challenge is to show ways in which evaluation can become an agent of transformational change of society, i.e., a *dynamic evaluation*, indicating characteristics that evaluations need to support transformational change. A third challenge is to identify the implications of dynamic evaluation for the development of evaluation capacities. These three challenges are considered in the following sections.

Definition of transformational change of society

There are different kinds of changes. Those that are *transformational* represent significant changes in direction and/or in size. In contrast with changes *a la Gattopardo* (i.e., changes so that everything remains the same), or micro-changes, transformational change of society makes a major and durable difference. In the context of the (long) march towards the SDGs, transformational change of society could amount to a significant advance in the achievement of one or more SDGs. Alternatively, it could involve a significant change in the direction in which society is organized[2] for a better achievement of the SDGs.

2 More recently, the relation between social transformation and the creation of a learning

Although the concern with transformational change is nowadays strong, half a century ago the importance of transformation for development was highlighted by Raúl Prebisch, the Latin American economist who created the United Nations Conference on Trade and Development (UNCTAD) and was the first Secretary-General of the Economic Commission for Latin America and the Caribbean (UN ECLAC – CEPAL).[3] In a 2018 blog, Zenda Ofir showed the importance of focusing evaluation on transformative development (Ofir 2018). Transformative evaluation has been presented as an evaluation paradigm by Donna Mertens (Mertens 2017; Mertens and Wilson 2018) at the community and individual level. In this essay the focus is at the level of society.

Evaluators should be aware of the possibility that the accumulation of small, incremental changes may lead to a transformational change of society, in analogy to what happened in some cases of industrial technical change. Evaluative evidence on this has been gathered in the ITAD evaluation of transformational change supported by the Climate Investment Funds. In its summing-up, the report states that "Incremental change represents a valuable contribution in progressing toward future transformation" (ITAD 2019, 47, para 134). However, although it is the case that incremental change may lead to transformation, this may not always be so, and there could be transformations without incremental change, as in the "Green Revolution".

In practical terms, evaluations that aim to be transformative would need to include not only the standard evaluation criteria (relevance, effectiveness/efficacy, efficiency, sustainability and impact), as well as *coherence* and also *significance*, which is similar to "materiality" (used by auditors).[4] An intervention (say a policy) may be highly relevant and get high marks on the other four criteria, and yet it may not change in any significant way the situation of the target population. Its actual effects, although positive, may be insignificant in terms of transformational change. As this cannot be captured through the "big five" criteria, it may be worthwhile to introduce another criterion, significance, that points to the size of the effect.

---

society has been discussed, and the importance of changing mindsets was considered to be at the root of success. See, for this purpose, Stiglitz and Greenwald (2014). It is worthwhile to mention that in the 40 pages word index of this book there is no reference to evaluation. This gap may be partly due to the rather narrow focus of evaluations (which would be different if evaluators change their mindset and practice, so that at least some evaluations become transformational, i.e., dynamic evaluations).

3 See Prebisch (1970).

4 Discussions with Robert Picciotto on the revision of the DAC criteria made me aware of the significance of significance as an evaluation criterion.

It should be noted that significance goes well beyond "statistical significance" since it also embraces the size of direct and indirect effects. It is concerned with "practical importance" (Ziliak 2008).

Dynamic evaluation

Whereas dynamic evaluations (DE) are by definition transformative at the level of society, none of the four types of evaluations generally considered (formative, summative, impact and developmental) are necessarily transformative at that level. As previously indicated, evaluations are expected to support learning and accountability, and discussions about the possible trade-offs between these two functions of evaluation are common. One side, frequently associated with NGOs, claims that accountability jeopardizes learning; the opposite view states that learning and accountability are complementary and synergistic. However, the transformative role of evaluation has not been an explicit part of this discussion. Evaluation has been excluded from the theory of change. The following paragraphs present different aspects of dynamic evaluations and may be able to show that dynamic evaluation can have a role in achieving transformation at the level of society.

This role of evaluation is consistent with Donald Campbell's evolutionary epistemology, more than with his view of the experimenting society (Campbell 1974), with evaluation contributing to the *social* (rather than natural) selection process. In the case of China's transformation, a sort of informal dynamic evaluation has been going on framed by Ang (2016) as "selection" without a direct reference to evaluation.

How can evaluation contribute to transformational change of society?

By *changing its focus* from projects and programs to strategies and policies, evaluation can become transformative, dynamic. It is not the case to forget the formers but to take them into account from the perspective of the latter. Also moving away from an emphasis on micro-issues, linearity (including the log-frame) and a single method towards macro issues, a complexity lens, and multiple methods. The needed shift is illustrated schematically in Table 1. The following paragraphs show practical ways by which such a shift in evaluations may be achieved.

TABLE 1. Shift in focus

FROM → TO	
Projects, programs	Strategies, Policies
Micro	Macro
Country	Global
Linearity	Complexity, Synergies
Single method	Multiple methods

SOURCE: AUTHOR'S OWN TABLE

A first exploration of key aspects of dynamic evaluations

A relevant evaluation agenda

The SDGs provide a menu of aspirations from which governments and civil society, at the national and subnational level, can choose what is relevant at their level. There are several key issues on which dynamic evaluations can and should help generating evidence and lessons learned on social innovation, incentives, inequity, climate change, migration, global public goods and bads[5], biodiversity, waste and plastics, health, and education for all. Evaluators can play the role of facilitators at an early stage, promoting a discussion between government and civil society/ Parliament at the national and subnational levels on the prioritization of themes for evaluations, taking into account the SDGs. Likewise, there are global issues for which Blue Marble Evaluation (Patton 2015) could be an appropriate way of dealing with them.[6]

Multiple methods and techniques

Evaluators should not be locked in single methods or single sources of data but should be willing to engage with different evaluation methods and techniques for capturing, processing, analysing, and synthesizing data. It is unlikely that all evaluators would be familiar with all relevant methods and techniques. Therefore, it would be essential to include in evaluation teams professionals with a variety of experiences and expertise, such as evaluation generalists (for example with expertise in systems approach), that may be able to coordinate dynamic evaluations. The openness and capacity to use

5 The literature on global public goods is much more abundant than that on global public bads. On the latter see, for example, Coyle and Ryan (2019) and Johansson and Kriström (2016), where "public bads" are presented as "negative externalities" (Johansson 2016, 21).

6 See also Picciotto (2014), and Clarke, Barnett and Van den Berg (2015).

different methods and techniques enhance the extent to which evaluation can make a contribution to transformational change of society.[7]

Big data and Information and Communication Technology (ICT) generated valuable opportunities for dynamic evaluations, as well as risks such as data overload, and the possibility of ending up with small ideas and masses of unused data. Artificial Intelligence (for example, through machine learning) could be of some help in identifying patterns. However, imagination, creativity, and sound judgment are the key to seize the opportunities and minimize the unintended negative consequences.

Scaling-up

The growing interest in scaling-up may be linked to the aspiration to achieve transformational change of society. Scaling-up is an important means to reach transformational change of society. Before discussing how evaluation can be used to support scaling-up and therefore transformational change of society, it is worthwhile to clarify a widespread view and misuse of evaluation in this respect: if a rigorous evaluation of a small scale intervention shows positive results this should not be considered as sufficient evidence that the intervention will also have positive results if it is scaled up, as there may be diseconomies of scale.

Moreover, if the rigorous evaluation of the small-scale intervention shows negative results, would this be an indication that the intervention should be scaled-up? Not necessarily. The argument that follows is presented in greater detail in Feinstein (2015). If there are economies of scale, an intervention that fails at a small scale may succeed after reaching a certain threshold. For example, an agricultural extension program may fail if it is directed towards a small population of 200 farmers in an isolated area without secondary roads and because of that extension workers have great difficulties in reaching these farmers. However, if the number of farmers would be scaled-up to 20,000, then the construction of rural roads may become feasible, and the program can be a success. A similar example would be one in which success at small scale does not lead to success at a higher scale due to the difficulty of staffing a large-scale program absent a sufficient supply of competent agriculturalists. In summary, neither success nor failure at small scale guarantees success when the intervention is scaled-up.

---

7 In terms of a self-referential or endogenous theory of change by which evaluation is expected to have effects on transformational change of society, the argument for mixed methods can be framed as increasing the probability that evaluation will have a positive effect on transformational change of society by expanding the means at its disposal.

How can evaluation support the scaling-up process and therefore transformational change of society? Can dynamic evaluations be used as a tool for effective scaling-up? Dynamic evaluations can help in identifying critical success factors for scaling-up and by conducting evaluations at different scales. Evaluation in the scaling-up context should be a dynamic evaluation, evaluating at different scales and at various points in time.[8]

Quasi-real-time evaluation

As it is well known, timely delivery and discussion of evaluations are crucial for their use, and this applies to all types of evaluations. In the case of dynamic evaluations, timeliness is particularly relevant because it may make a difference in terms of being able to influence a social transformation. Development evaluators could learn from the practice of real-time evaluation (which should not be confused with monitoring) that is used in evaluations of humanitarian aid. Useful references here are Polastro (2014) and Feinstein and Beck (2006).

New opportunities for quasi-real time evaluations are provided by big data, which is now being used. It is likely that it will play a growing role in policy evaluation, especially since results may be delivered almost in real-time. So, evaluators must develop their competencies to work with it, and collaborate with professionals already experienced in using big data (see Højlund et al. 2017).

By proceeding in this way, evaluators practicing dynamic evaluation will be well positioned to timely support policy makers in decision-making that can contribute to social transformations taking into account results from the actual implementation of policies. Thus, formative evaluations can become transformative at the societal level.

Politically sensitive evaluation

Dynamic evaluations could enrich the political debate, strengthening deliberative democracy by providing evidence and evidence-based arguments on key questions related to crucial themes such as health, education, infrastructure. Given the strong populist trend in the world, dynamic evaluations must pay particular attention to the sustainability of results, showing whenever it is the case that successful short-term results do not guarantee long term results. In fact, they may even jeopardize

8 Parrot and Carman's recent article on this issue focuses only on process evaluations (Parrot and Carman 2019). For a more general discussion on scaling-up and external validity, with specific examples, see Feinstein (2017). A recent comprehensive and promising approach to evaluation for scaling-up is provided in McLean and Gargani (2019).

the achievement of sustainable results when resources are used as hand-outs to get political votes.

To navigate the troubled waters of an exceptionally adverse political climate and considerable differences in opinion on the roads to follow towards systemic changes, it is worthwhile to consider an evaluation approach that, as Robert Picciotto (2019) recently pointed out, is now nearly forgotten.

This approach can nurture the democratic debate, strengthening deliberative democracy, which would be particularly timely given the current polarization of most societies. What is crucial is to consider different perspectives, and this could be done even within a single evaluation, focusing first on achievements (intended and unintended), then on shortcomings (or failures or limitations), and finally making a judgment taking into account the achievements and shortcomings. A practical way to proceed is ensuring that the scope of work (or terms of reference) directs the attention of the evaluator(s) to both achievements and shortcomings. This more modest approach to adversary evaluation may be more feasible than a full-fledged adversary evaluation. It should be noted that the thrust of Picciotto's article is indeed that the judicial evaluation model is usually too demanding and that it is only one embodiment of adversary evaluation. According to Robert Picciotto (personal communication), simpler approaches are in fact feasible and often preferable.

Nevertheless, adversary evaluation could be reframed in terms of a dialectical approach with three phases: "thesis", "antithesis" and "synthesis". Note that by focusing on both positive and negative results (the positive results presented as the "thesis", the negative as the "antithesis"), the approach avoids two rather frequent types of biases: the "positive bias", which neglects negative results (or shortcomings) and the "negative bias", which neglects positive results (or achievements). Thus, it can be perceived as an approach that promotes impartiality (through a kind of explicit and compensated partialities).

These two first phases could be seen as "deconstructing" or decomposing the evaluated intervention in terms of positive and negative results, which correspond to benefits and costs. In these two phases, assessments should be made of the worth and merit of the intervention, taking into account its relevance and sustainability (two criteria associated with worth, "doing the right things") as well as its effectiveness (evaluation criterion associated with merit, "doing things right").

A third phase, "synthesis" (which is neither included in adversary evaluation nor in Picciotto's recent article (Piccioto 2019)) would be fully evaluative, judging the

value of the intervention in light of the evidence and judgements provided by the two first phases, focusing on the impact of the intervention, taking into account its efficiency. The synthesis is a phase of the evaluation in which creative thinking should play an important role, "putting the data together in new ways to see the interactions among separate findings more holistically; synthesizing diverse themes in a search for coherence and essence" (Patton 2018, 22).

Making context matter

For an adequate consideration of context in conducting dynamic evaluations and their synthesis, it is worthwhile to use an adaptation of "realistic (or realist) evaluation" based on the triad *Context*, *Interventions*, and *Results*. It can be used to systematize the knowledge gained from the evaluations of interventions of different types, to identify which interventions achieved positive or negative results in specific contexts. Patterns of success and failures may be thus identified for different contexts and interventions which can be of use in determining levers for transformational change, taking into account specific national and subnational contexts.

Complexity is an aspect of the context that matters for the evaluation of SDGs interventions. An unintended consequence of the SDGs is that each SDG may end up being considered as an entity in itself, pursuing it in isolation from the other SDGs, as if the "S" corresponding to "Sustainable" would be for "Silo". The intersectoriality of the SDGs and their synergies should be given due importance during their implementation and when evaluation takes place (see, for example, Alcamo et al. 2018 and Tett 2015).

Structurally linear tools, such as the log frames, should be replaced by an evaluation framework that takes complexity into account, considering the SDGs as components of a system, and yet it is simple enough to facilitate its use and the communication of results. An adaptation of the Cynefin framework (Snowden and Boone 2007) may be appropriate and could be combined in some dynamic evaluations with the Context, Interventions, and Results triad.

Third loop learning and dynamic evaluations

Triple-loop learning is defined as transformative learning (collectively examining underlying assumptions, leading to change in attitudes and social norms) in contrast with single-loop learning – instrumental learning (acquiring new knowledge individually, without assessing implicit assumptions) and with what has been called double-loop communicative learning (understanding/interpreting knowledge through interaction with others) (Mockbee and Newsham 2013). It has been acknowledged that

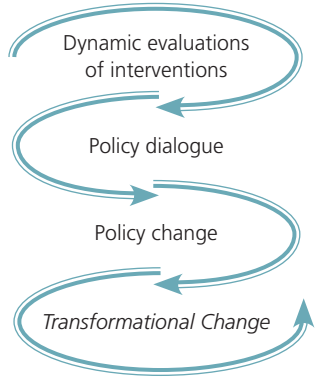
> Triple-loop learning is a slow process. It requires us to question our own assumptions, how these affect others and our ways of working, to be able to gauge the extent to which transformative change is necessary, and what it would look like. At the same time, it also requires us to build the relationships that would bring about the collective behaviour that is a precondition for transformative change (Mockbee 2013, 24-5).

Therefore, it is important to manage expectations concerning transformational learning, given the requirements for triple-loop learning. Dynamic evaluations can contribute to triple-loop learning with an examination of assumptions, including those concerning values (Nkwake 2019, forthcoming), and it can also use findings from behavioural economics (World Bank 2015).

Evaluation of policy dialogue as an instrument for transformational change of society and institutional development

The shift of focus from projects and programs does not imply that they should no longer be considered as development interventions. However, a transformational evaluation has to take into account the extent to which they are contributing to transformational change and institutional development, for example by opening windows for policy dialogue that may induce policy change and through it transformational change may take place.[9] Figure 1 may help to see the main direction of causality.

FIGURE 1. Main directionality of causality.



9 For the evaluation of policy dialogue see IFAD (2015), and for policy change and its political economy, Corduneanu-Huei et al. (2013).

Inequalities, innovation, transformational change of society and dynamic evaluations

The development experience of China since the late 1970s is perhaps the most extraordinary case of transformational change of society. It is a change that consisted of economic growth during decades at growth rates of approximately 10%, with an extraordinary poverty reduction effect but, at the same time, with an increase in inequalities. One of the ways in which dynamic evaluation can and should play a role in these processes of transformational change of society is by identifying innovations that reduce (or have a potential for reducing) inequities. That is, inequities-reducing-innovations (a sub-set of pro-poor innovations) could pave the way for a transformation that mitigates or eliminates increased inequities.

Capacities for the practice of dynamic evaluation

Are there any challenges and/or implications of dynamic evaluations in terms of evaluation capacities? To address this question, it is useful to disaggregate evaluation capacities in the following way:

a. Capacity to conduct evaluations
b. Capacity to manage evaluations
c. Capacity to demand evaluations

The capacity to conduct dynamic evaluations requires awareness of complexity and tools to deal with it, as well as focused attention on significant changes. Relevant materials to develop the capacity to conduct dynamic evaluations have been made available as, for example, by Stephens, Lewis and Reddy (2018).

To make significant progress in the development of capacities to conduct dynamic evaluation it would be worthwhile to promote a massive involvement of universities in evaluation, transforming a potential into an actual supply of dynamic evaluations.

Rather than a sporadic participation of universities in evaluations, which is the "business as usual" approach, for dynamic evaluations it would be worthwhile to ensure a massive systematic involvement of universities at the national and sub-national levels. This engagement could be done through their direct participation in evaluation teams, and their complementary contribution, preparing background papers and surveys, as well as theses on subjects potentially relevant for dynamic evaluations. Proceeding in this way may not only enhance the quality of evaluations

but would also be a means to promote the development of evaluation capacities through "learning by doing". In addition, it would help to establish and/or enhance links between research and evaluation. Through their participation in evaluations researchers may be able to identify topics for research that may prove to be useful in the design of interventions that contribute to transformational change, and which could also include evidence gathered by evaluations.

Thus, universities would be generating a stock of data, information, and knowledge, as well as evaluation capacities that like "dormant cells" would be ready to enter into the evaluation battlefield. Donors and international organizations, including United Nations (UN) agencies, Multilateral Development Banks (MDBs) and the European Commission (EC), who are evaluation partners committed to the SDGs, should be encouraged to involve national and subnational universities in their SDGs evaluations.

A key challenge in the capacities to manage evaluations is ensuring that the evaluations' terms of reference direct the attention of evaluation teams to identify and assess significant changes and to do so taking into account synergies and unintended effects. One implication is that the capacity to manage evaluations should include the skill to write adequate terms of reference for dynamic evaluations.

Furthermore, it is important to develop the capacity to demand dynamic evaluations. It requires that society's decision-makers in government and civil society become aware of the role that evaluation can play in achieving transformational change of their societies.

Finally, the IDEAS proposal that will be discussed at the 2019 Global Assembly, which is to create an International Evaluation Academy, could help to identify systematically the capacities required for dynamic evaluation. This Academy can link evaluation with the efforts universities and sciences are doing to understand in a multi- and transdisciplinary way how societies, economies, and the world function and can be transformed to achieve the SDGs.

References

Alcamo, Joseph, Caroline Grundy, and Jorn P.W. Scharlemann. 2018. *Interactions among the sustainable development goals and why they are important.* Brighton: SSRP (Sussex Sustainability Research Programme), University of Sussex and Institute of Development Studies (Policy Brief 1). https://tinyurl.com/y33eycjh.

Ang, Yuen Yuen. 2016. *How China escaped the poverty trap.* Cornell University Press: Ithaca.

Campbell, Donald T. 1974. Evolutionary epistemology. In: *The philosophy of Karl Popper*, vol 1, (ed.) Paul A. Schlipp, 413–463. LaSalle IL: Open Court.

Clarke, Charlotte, Chris Barnett, and Rob van den Berg. 2015. *Report New frontiers for evaluation in an era of market-oriented development.* Steyning: Wilton Park. https://tinyurl.com/q97m2w7.

Corduneanu-Huci, Cristina, Alexander Hamilton, and Ferrer Issels Masses. 2013. *Understanding policy change. How to apply political economy concepts in practice.* Washington D.C.: The World Bank. https://tinyurl.com/y375nmct.

Coyne Cristopher J., and Ryan E. Matt. 2019. Foreign intervention and global public bads. In: *Economic and Political Institutions and Development*, (eds) Hall Joshua and Sarah Harper. 41-59. Springer: Cham.

Feinstein, Osvaldo. 2015. Getting to scale. How to bring development solutions to millions of poor people – Book Review. *Development in Practice* 25 (2): 289-290.

Feinstein, Osvaldo. 2017. Trends in development evaluation and implications for knowledge management. *Knowledge Management for Development Journal* 13 (1): 31-38.

Feinstein, Osvaldo, and Tony Beck. 2006. Evaluation of development interventions and humanitarian action. In: *The SAGE Handbook of Evaluation*, (eds) Ian Shaw, Jennifer Greene, and Melvin M. Mark, 536-558. Thousand Oaks, CA.: SAGE.

Højlund, Steven, Karol Olejniczak, Gustav Jakob Petersson, et al. 2017. The current use of big data in evaluation. In: *Cyber society, big data, and evaluation*, (eds) Gustav Jakob Petersson, and Jonathan D. Breul, 35-60. New York: Routledge, New York. (Comparative Policy Evaluation 24).

International Fund for Agricultural Development (IFAD). 2015. *Evaluation manual*. https://tinyurl.com/yxwmnted.

ITAD. 2019. *Evaluation of transformational change in the Climate Investment Funds.* https://tinyurl.com/y2dcdsan.

Johansson, Per-Olov, and Bengt Kriström. 2016. *Cost-benefit analysis for project appraisal.* Cambridge: Cambridge University Press.

McLean, Robert, and John Gargani. 2019. *Scaling impact. Innovation for the public good.* New York: Routledge.

Mertens, Donna M. 2017. Transformative research. Personal and societal. *International Journal for Transformative Research* 4 (1). https://tinyurl.com/y5zkq48k.

Mertens, Donna M., and Amy T. Wilson. 2018. *Program evaluation theory and practice.* 2nd ed. New York: Guilford Press.

Mockbee, Michelle, and Andrew Newsham. 2013. *Acting on what we know and learn for climate development policy. Final report.* Brighton: Institute of Development Studies. https://tinyurl.com/yy9vqnqh.

Nkwake, Apollo M. 2019. *Working with assumptions in international development program evaluation.* 2nd ed. New York: Springer (in press).

Ofir, Zenda. 2018. *Evaluation for transformation 1. Can we please get beyond 'buzzwords'?* https://tinyurl.com/y24hdp9v.

Parrot, Adria, and Joanne Genova Carman. 2019. Scaling up programs. Reflections on the importance of process. *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme* 334 (1): 131-138, Spring 2019/Printemps 2019. https://tinyurl.com/y5o7c2d6.

Patton, Michael Quinn. 2015. *Blue marble evaluation.* https://tinyurl.com/y47uajn8.

Patton, Michael Quinn. 2018. A historical perspective on the evolution of evaluative thinking. *New Directions for Evaluation* 2018 (158): 11-28. https://tinyurl.com/y6ckoke5.

Picciotto, Robert. 2014. The future of development evaluation. *Revista de Evaluación de Programas y Políticas Públicas* 2014 (3): 259-287. https://tinyurl.com/y2s58u6m.

Picciotto, Robert. 2019. Is adversary evaluation worth a second look? *American Journal of Evaluation* 40 (1): 92-103.

Polastro, Riccardo. 2014. Evaluating humanitarian action in real time. Recent practices, challenges, and innovations. *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme* 29 (1): 118–134. https://tinyurl.com/y546sw8z.

Prebisch, Raúl. 1970. *Transformación y Desarrollo.* México D.F.: Fondo de Cultura Económica.

Snowden, David J., and Mary E. Boone. 2007. A leader's framework for decision making. *Harvard Business Review* 85 (11): 68-76, 149. https://hbr.org/2007/11/a-leaders-framework-for-decision-making.

Stephens, Anne, Ellen D. Lewis, and Shravanti Reddy. 2018. *Inclusive systemic evaluation for gender equality, environments and marginalized voices (ISE4GEMs). A new approach for the SDG era.* New York: United Nations Entity for Gender Equality and the Empowerment of Women (UN Women). (UN Women evaluation guidance series). https://tinyurl.com/y4dyps6o.

Stiglitz, Joseph, and Bruce Greenwald. 2014. *Creating a learning society. A new approach to growth, development, and social progress.* New York: Columbia University Press.

Tett, Gillian. 2015. *The silo effect.* New York: Simon & Schuster.

World Bank. 2015. *World development report 2015. Mind, society, and behavior.* https://tinyurl.com/o99s68z.

Ziliak, Stephen T., and Deidre N. McCloskey. 2008. *The cult of statistical significance.How the standard error costs jobs, justice and lives.* Ann Arbor: University of Michigan Press.

Revolution in the making: evaluation "done well" in the era of the SDGs with a youth participatory approach

Bianca Montrosse-Moorhead, Khalil Bitar, Josette Arévalo, and Antonina Rishko-Porcescu

ABSTRACT. Evaluation done well is and always has been transformational — revolutionary even. The markers of these revolutions include the convergence of incremental, reform-oriented, global, and transformational changes. A new evaluation revolution fueled by youth is in the making. We explain this new revolution in light of the significant, large-scale challenges we face as a globe, and the proposed solutions articulated in the 2030 Agenda for Sustainable Development. In this chapter, we lay out the enabling conditions that are giving rise to a youth-participatory evaluation revolution. We then describe the architectural outlines of youth-participatory evaluation, including what makes it distinct, who some of the critical champions are, and the value-added of this revolution. Collectively, this chapter offers a novel answer to a fundamental evaluation question — what does "done well" mean and what ought it look like in practice? This chapter ends with a call to action for the evaluation community.

Introduction

Evaluation done well is and always has been transformational. These transformations are not solely marked by incremental changes guided by doing more of the same, just better, to improve performance. Nor are they solely marked by reform-oriented changes, which are focused on rules, structures and processes in order to make changes to a system and/or its constituent parts. Transformational change in evaluation has always led to innovation and the creation of unimagined possibilities (SDG Transformations Forum).

One example from our history is when use became a focus and significantly changed every aspect of our work (Patton 1978). This shift was seismic in the sense that it not only broadened our understanding of the purposes of evaluation, it also shifted the field's understanding of the way evaluation can work, shifted power structures, and led to the development of new techniques and tools. It also spawned an entire line of research on evaluation utilization.

A more recent example of transformational change in evaluation is the mainstreaming of social justice approaches to evaluation (Mertens 1999). This shift was marked by some of the same ground-breaking changes, for example, broadening understanding of the purposes of evaluation and expanding conceptions about who does versus who ought to have power. It was also marked by the integration and mainstreaming of a new logic in evaluation, namely, the philosophically grounded transformative paradigm, which has distinct answers to the nature of ethics, of reality, of knowledge, and of systematic approaches for knowledge production (Mertens and Wilson 2012). These are only two examples of transformational change — Kuhnian revolutions[1] even — in the evaluation field, but the evaluation profession is rich with many more examples. We also believe it is time for a new Kuhnian revolution.

In this chapter, we return to a fundamental question in evaluation – what does "done well" mean and what should it look like in practice. Here, we first make the case that a Kuhnian evaluation revolution is in the making and is being fuelled by three enabling conditions: a) the 2030 Agenda for Sustainable Development (United

1 Thomas S. Kuhn, an American philosopher of science, introduced the notion of scientific revolutions in his 1962 book *The Structure of Scientific Revolutions*. His central thesis was that the development of science is marked by multiple periods of stability and transformation and that these revolutions completely revise existing scientific beliefs, practices, or both.

Nations Global Assembly 2015); b) the state of the world; and c) youth and young people themselves. We then describe how, both within and outside of evaluation, youth have not had a place at the table as equal partners. We then make a case for the evaluation revolution we are calling for – genuine youth-participatory evaluation. In doing so, we touch upon what youth-participatory evaluation is and how it is distinct from other approaches to evaluation, who some of the champions of this approach within the evaluation landscape are, and what the value-added is of having youth at the table as equal partners, including young and emerging evaluators. Collectively, this chapter offers a redefinition of what "evaluation done well" means, including implications for evaluation practice. This chapter is a call to action for the evaluation community. The revolution is coming. What role will you play in its history?

The Sustainable Development Goals (SDGs), youth, and evaluation: three necessary ingredients for revolution

The time to act is now

The world is not burning, but the embers are lit. As the 2030 Agenda for Sustainable Development makes clear, the time to act to put out this fire is now. This fire is being sparked by several significant, large-scale challenges that are not country specific. These challenges are world-wide and affect all of us. Some of these challenges include billions of citizens still living in poverty (Lattimer et al. 2018); rising inequalities within and among countries (UNDP 2018); disparity of opportunity, wealth, and power especially prevalent among girls and women and youth (World Economic Forum 2018); violent conflicts and related humanitarian crises displacing millions of people (Lattimer et al. 2018[2]; UN Women 2019); and climate change (United Nations 2015), to name a few.

The shared blueprint for action, and a cause for hope, the 2030 Agenda for Sustainable Development lays out 17 Sustainable Development Goals (see Figure 1). Collectively, these 17 goals and associated targets coalesce around putting people, planet, prosperity, and peace first. Moreover, it recognizes that the challenges we face as a globe, and the solutions they will require must be done in partnership.

2 According to Lattimer et al. (2018, 19), "The total number of people forcibly displaced due to conflict, violence or persecution reached 68.5 million in 2017, an increase of 2.9 million (4.5%) from 2016, the sixth consecutive annual increase."

In short, the only way to the type of revolutionary change needed is together and different. We must do things together, as many of the challenges are not country-specific, as is the case for climate change. At the same time, we must do things differently – it cannot be more of the same with the same actors. This is why, for example, the private sector was included as both a partner and co-responsible development actor; a first in a global development agenda.

FIGURE 1. The 17 SDGs reproduced with permission from the United Nations.

AVAILABLE FROM: https://www.un.org/sustainabledevelopment/news/communications-material/

Youth must have a seat at the table as equal partners to realize revolutionary change

There are no SDGs devoted explicitly to youth. Only SGDs 4 (quality education), 8 (decent work and economic growth), and 13 (climate action) specifically call out the youth in their targets. And, still, none focus exclusively on youth.

There are three ways to interpret this. One is that other groups have more pressing needs than youth, which is why no SDG is devoted to youth exclusively. There is little evidence to support this interpretation. The United Nations General Assembly Resolution which provided the backbone for the 2030 agenda states:

> People who are vulnerable must be empowered. Those whose needs are reflected in the Agenda include all children, youth, persons with disabilities (of whom more than 80 percent live in poverty), people living with HIV/AIDS, older persons, indigenous peoples, refugees, and internally displaced persons and migrants. (United Nations Global Assembly 2015, para 23)

Moreover, 20 targets across six SDGs (Zero Hunger, Quality Education, Gender Equality, Decent Work and Economic Growth, Reduced Inequalities, and Climate Action) have a strong child or youth focus, and just over one third of SDG targets explicitly or implicitly refer to children and youth (UNDP 2017).[3] A second is that youth were not considered because they were not consulted during the process of writing the SDGs. It is unclear if there is evidence for this position. A third is a common dilemma in these matters – is it better to integrate youth into all aspects, or is it better to call attention to their exclusive perspective through, for example, creating a "Youth Equality" SDG? It is unclear whether and to what extent deliberations occurred regarding giving youth their own SDG. What is clear is that the UN has had a youth focus for some time (United Nations General Assembly. 1995). This is because there are several pressing issues affecting this group directly including, for example, that despite an increase in absolute numbers, the proportion of young people in the world is dwindling; educational opportunities are not guaranteed; and health is a common concern, among others.[4] From a United Nations perspective, there is little evidence to suggest that other age groups would have more pressing needs.

Regardless of the reasons why and having recognized that children and youth bear the brunt of what we do or do not do now, several UN agencies have set out to ensure youth are at the table. The view of these agencies is that youth and youth issues are cross-cutting or embedded in the conceptualization of all SDGs, thus giving them their own SDG made no sense. For example, the UN Youth Strategy 2030 (2018, 5) emphasized the role of youth in achieving Agenda 2030:

> The [UN Youth] Strategy aims to facilitate increased impact and expanded global, regional and country-level action to address the needs, build the agency and advance the rights of young people in all their diversity around the world, and to ensure their engagement

3 We adopt the United Nations definitions of these terms (United Nations General Assembly A/36/215 1981). Children are persons under the age of 14. Youth are persons between the ages of 15 and 24 years and can be further divided into teenagers (15-19) and young adults (20-24).
4 Compiled from various UN documents.

> and participation in the implementation, review and follow-up of the 2030 Agenda for Sustainable Development as well as other relevant global agendas and frameworks.

The recently adopted Lisboa+21 Declaration on Youth Policies and Programmes (2019) echoes the value of youth and their role in the SDGs, "the Assembly recognize[s] children and youth as agents of change and recognize[s] that the Sustainable Development Goals are integrated, indivisible and universal in nature, and therefore that all of them apply to youth" (p. 1).

Having youth at the table is more paramount than ever. Currently, we have the largest generation of young people in human history – 1.3 billion aged between 10 and 24 to be exact, which is roughly 30 percent of the world's population[5]. And, around 90 percent of the 30 percent are located in the developing world, meaning revolution cannot happen without them at the table. In fact, as United Nations Secretary-General António Guterres has been arguing for some time, "it's not enough to listen to young people and provide a seat at the table – we need to take a seat at your table".[6]

It is not only the sheer number of youth that makes a case for them to be at the table, or for us to sit at theirs, it is also the environment. UNICEF's *Unless We Act Now* (2015) report makes a convincing case that children and youth bear the ramifications of what we do and do not do regarding climate change. While it is true that climate change affects some groups of children, youth, and their families disproportionately – specifically those living in high flood zones, high drought zones, or high poverty zones – it is also true that it affects all of us. After all, climate change is not restricted to national, geopolitical, or high need boundaries.

The coming revolution: moving from evaluation on youth to evaluation with and by youth

The adoption of the 2030 Agenda has implications for evaluation. It puts an increased emphasis on country-led evaluations and building capacity through strengthened data systems and evaluation programs with each country. These country-led evaluations take place as Voluntary National Reviews (VNRs). They have several aims that include monitoring the progress of implementation of the 2030 Agenda and sharing experiences, successes, challenges, and lessons learned.

5 Compiled by authors from UN and CIA sources.
6 UN News 23 June 2019, https://news.un.org/en/story/2019/06/1041111

Since the first VNRs were conducted, several challenges have surfaced, including those that affect youth. Despite growing recognition that youth are cross-cutting or embedded in all SDGs, to date, very few published VNRs have included youth in the creation of reports.[7] Of those that have, most have included youth voices by only collecting data from them. Moreover, as Patton (2017, xvii) noted in his review of a five-year strategic evaluation plan of a major international agency, "intended beneficiaries of development efforts were essentially invisible. The evaluation appeared people-less and heart-less", especially when it came to youth.

To the extent that we take seriously the idea to leave no one behind, and to the extent that evaluation norms and standards uphold a do-no-harm normative view, we are in danger of violating both if evaluators across the globe and those who commission evaluations do not fully embrace a rights-based approach to evaluation. A rights-based approach means you put people and heart first.

There is a reason to be hopeful. Recent literature has put equity-focused and gender-responsive evaluations at the forefront (Segone and Rugh 2013; Segone and Tateossian 2017). We have also seen the rise and integration of culturally responsive evaluation and indigenous perspectives in evaluation (Cram, Tibbetts and LaFrance 2018; Hood, Hopson and Frierson 2005). Moreover, within the United Nations system, several positive strategies have recently been adopted. To facilitate the integration of youth perspectives in the implementation of the SDGs, the United Nations launched a system-wide Youth Strategy in 2018 (Youth 2030). The strategy aims to step up efforts to address the needs, amplify the voice and advance the rights of young people and calls for their meaningful engagement and participation in the implementation, review, and follow-up of the 2030 Agenda. Each United Nations agency has also taken steps to put youth at the centre of the 2030 Agenda. The 2019-20 UNFPA Adolescent and Youth Strategy *(forthcoming)*, for example, articulates a framework of prioritization for the delivery of UNFPA's mandate that puts the rights of young people to make informed choices over their body, their life, and their world at the centre. This is strengthened by their 2019 Revised Evaluation Policy that, for the first time, makes explicit reference to the inclusion of young people as key stakeholders to ensure useful and credible evaluation results. UNICEF also recently released a guidance note on adolescent participation in UNICEF monitoring and evaluation (UNICEF 2018). The next step, given that all roads lead to youth, is to put them at the centre through youth-participatory evaluation.

7 See the Voluntary National Reviews Database at https://sustainabledevelopment.un.org/vnrs/

"Youth-participatory evaluation" is defined as the process of involving young people in conducting evaluations (Checkoway and Richards-Schuster 2005). It includes the participation of young people in different roles (evaluation managers, evaluators, evaluation reference group members, and informants) in all phases of an evaluation (preparatory phase, design phase, data collection phase, analysis and reporting phase, and facilitation of use and dissemination phase). The purpose of the youth-participatory evaluation is to empower young people, to recognize their potential, and to acknowledge their legitimate and unique perspectives by meaningfully and sustainably engaging them in evaluation and by focusing on issues that affect their lives. In short, it moves evaluation from being framed as something that is done *to* young people to evaluation *with* or *by* young people.

Despite the fact that youth-participatory evaluation has not been mainstreamed in evaluation practice, a body of literature is available from which to draw. Literature exists that establishes the benefits of youth participation in evaluation (Laws and Gillian 2004; London, Zimmerman and Erbstein 2003; Purdue, Peterson and Deng 2018; UNICEF 2018; Zeldin, Bestul and Powers 2012). The majority of this literature distinguishes between benefits for evaluation quality, young people as individuals, the communities in which youth reside, and organizations that conduct or commission youth-participatory evaluations. A smaller body of the literature discusses barriers and challenges to youth participation in monitoring and evaluation activities (Camino 2005; Hulshof 2019). The literature, in general, identifies four different modalities for participation that are linked to the different functions in which youth can participate in evaluations (Checkoway and Richards-Schuster 2003):

- Young people as informants for an evaluation *(The business as usual approach)*
- Young people as consultants/advisors for evaluation *(The sometimes approach)*
- Young people as co-evaluators in an evaluation *(The rarely approach)*
- Young people as leaders/directors of an evaluation *(The almost never approach)*

While a comprehensive review of each of these modalities and underlying assumptions is beyond the scope of this chapter[8], generally speaking, the business as usual approach is to include young people as informants for an evaluation. In this

8 To learn more see Checkoway and Richards-Schuster (2003) and Richards-Schuster and Elliott (2019).

scenario, youth are consulted and give feedback, but their engagement is limited, and they generally have no power in the decision-making process. Sometimes, youth have been positioned as consultants or advisors in an evaluation. In this scenario, youth are somewhat engaged, but in very specific or limited ways. For example, they may work with adult evaluators to support data collection and have power within that specific area, but do not contribute to decision-making in other aspects of the evaluation. By and large, the rarely used approaches in evaluation are those where youth are positioned as co-evaluators or as leaders/directors. In the co-evaluator scenario, youth and adult evaluators work as a team to envision, design and implement all aspects of an evaluation. As such, youth are very engaged and share decision-making power with adults throughout the evaluation process. In the leader or director role, it is youth, not adults, who are in charge. Thus, youth envision, design, and implement all aspects of an evaluation and hold all of the decision-making power, including the framing and development of evaluation questions that are important from a youth perspective. Adult evaluators serve as coaches or facilitators.

The revolutionary potential of youth participation resides exclusively in the modalities that position youth as co-evaluators or as leaders or directors of the evaluation. Said another way, the Kuhnian revolution we are calling for involves youth as co-evaluators or leaders. Moreover, practically relevant information on youth-participatory evaluation is available across several different resources (Checkoway and Richards-Schuster 2003; Cousins (in press); Richards-Schuster and Elliott 2019; UNICEF 2018; Wridt 2018). These resources cover the ethics of youth participation, the nature of reality among this group and how different it is from business as usual, and what it means to produce knowledge with and by youth. In this way, the barriers to mainstreaming youth-participatory evaluation have been addressed, and the path to revolution cleared.

What value added do youth and YEEs bring to the table in this landscape

A genuine and effective revolutionary change – especially within the SDGs framework – needs youth and youth issues to be re-considered as we have argued above. Youth inclusion and consideration in the implementation of all targets of all 17 SDGs, both on the country and global levels, enhances the potential of reaching these targets and the SDGs and to achieving a better and more sustainable future for all per the Agenda 2030 language. Young people overall have several overwhelmingly common characteristics that, we argue, highly enhance the

potential of successfully achieving these noble global objectives and goals. These characteristics, while common, speak about general trends, so they should not be understood in any way to apply to all youth as youth are not monolithic. There are several rationales to consider, but we focus on three here.

First: values youth have and bring

Youth around the world are sparking, leading, and taking a highly active part in multiple socio-political grassroots movements advocating for enhanced social justice, equality, equity, environmental sustainability, peace, and several other rights-based and forward-looking critical issues across the globe. Such movements are national, regional, and international, but many of them are effectively global in nature or quickly become so given the shared challenges our world is experiencing and the outstanding qualities of youth to get their message across national borders, quickly. While traditional political parties in many industrial and developing nations alike are retracting while challenged by strong "populist" movements and political parties, youth in many of these countries are leading efforts to counterbalance such extreme political tendencies. They are fighting for more, not less, social justice, openness, and environmental protection actions and policies. To illustrate, in the last European Parliament Elections in 2019, the Green Party alliance posted its strongest ever performance in these elections winning 69 seats. Such results are said to have been highly influenced by youth marches calling for political action over climate change, for example, #FridaysForFuture, including the United Kingdom, Ireland, France, and Germany, where the Green Party alliance did particularly well (Guy and Regan 2018). In the same vein, youth activists across the globe are promoting the SDGs and calling for a stronger consideration of them in the national plans and policies of their countries and across regions including movements such as *Be The Change*, *LittlexLittle*, and *Not too Young to Run.*[9]

As said above, however, we recognize that youth are not a homogeneous group where all its members agree on the same principles. In each country and region – and indeed on the global level – youth, as a group, carry political views that extend from the extreme right to the extreme left of the political spectrum. Youth are also influenced by a host of political, economic, and social factors that influence all other members and sections of the rest of the society. Different members of the youth of any given country or region respond to, and interact with, these factors in different ways.

9 See https://www.un.org/sustainabledevelopment/youth/

Yet, what we are talking about here — and what we want to focus attention on — are strong trends of youth activism and the issues leaders and activists are often working on: socio-political grassroots movements advocating for enhanced social justice, equality, equity, environmental sustainability, peace, and several other rights-based and forward-looking critical issues across the globe. For example, while it is true that a recent poll in the United Kingdom found that a third of young voters now believe the army should run the country instead of the parliament (Gray 2019), the understanding of these views must be contextualized. While the results do show a minority of youth to have such views, much of these views can be attributed to the political landscape in the United Kingdom after Brexit and the frustration the overall population feels after the failure to reach a final deal on Britain's exit from the European Union with strong political division in the British Parliament on the issue. After all, on the Brexit vote, when 51.9% voted to leave the European Union and 48.1% to remain, over 70% of 18 to 24-year-olds voted to remain, and under 30% opted to vote to leave the European Union. This is in contrast to the 40% of those aged 65 and over who supported remaining. One of the main problems then was that only about 64% of 18 to 24-year-olds voted, while 90% of over-65s turned out to vote (Spratt 2018). Such trends, however, are changing and we see increased participation of youth in elections as was observed during the European Parliament Elections in 2019 referenced above.

Young evaluators carry – and act upon – these very values and principles. They are often also activists or leaders in many of these social movements themselves. Even when not directly linked to such movements, young evaluators engage with them through social media outlets, which many of these social movements capitalize very effectively. Young evaluators are also building stronger networks and alliances with other equity and social justice-focused groups. For example, EvalYouth – an EvalPartners youth-led global network that promotes young evaluators to become leaders in the evaluation field – has established strong connections with EvalGender+ and EvalIndigenous, in addition to EvalSDGs, the Global Parliamentarian Forum among other similar groups and networks. Moreover, in carrying out this work, young evaluators are not methodologically dogmatic. More often than not, they tend to embrace mixed methods, contextually responsive approaches, and understand that rigor is not the property of technical methods, but rather of evaluative thinking.

Second: skills youth have

Youth, including young evaluators, have, on average, a robust skillset that is highly relevant and useful for genuine and effective revolutionary change. Such

skills include, but are not limited to, young people's technological savviness and connectivity.[10] Any change, especially revolutionary changes, in this case, needs to be communicated effectively. With their technological savviness and technological advances, youth are communicating faster than any previous generation. Social media outlets are being utilized highly effectively by young activists to promote and advance several of the social movements discussed above.

Young evaluators across the globe are cooperating in unprecedented ways. For example, EvalYouth has established a free, multilingual eLearning agenda to democratize evaluation knowledge and share successes and lessons learned. Several new initiatives are led by young evaluators to enhance the implementation of the SDGs and their countries' Voluntary National Reviews. Such initiatives are calling for stronger youth participation and consideration of social justice and environmental sustainability issues that concern young people. What is most important about these initiatives is that young evaluators are coordinating them in multiple countries at the same time. They are considering all issues affecting youth and other members of their society with the "youth element" being a cross-cutting one that is relevant to all SDGs and issues young evaluators are generally promoting.

Third: youth supply

This third element might be more apparent even for young evaluators specifically. While in previous decades the evaluation field was practiced by a limited number of evaluators especially from North America and Western Europe, this trend has significantly shifted in recent years. Many developing countries, which only had a handful of evaluators during the past decades, are seeing a significant increase in the number of young and emerging monitoring and evaluation professionals. A prominent trend of supply of young evaluators is especially observed in Sub-Saharan Africa, the Middle East and North Africa (MENA), and South America. Young evaluators in these regions are also highly eager to build and enhance their evaluation skills. Over 50% of all applications received by EvalYouth for its flagship program – the Global Mentoring Program – in its two phases so far came from young evaluators and monitoring and evaluation professionals from the African

10 Derived from a poster presentation of Josette Arévalo, Antonella Guidoccio and Claudia Olavarria on behalf of EvalYouth LAC, at the annual American Evaluation Association Conference in Atlanta, Georgia, USA in 2016.

continent alone. These young and emerging evaluators often provide a fresh perspective dealing with many of the issues in hand and their countries' national priorities and in relation to the SDGs. This surge in supply of evaluators is a positive development as specialists are needed for extended, more regular, and quality monitoring and evaluation of the SDGs in all countries.

Who is doing what to change the status quo?

Understanding the critical role of youth and the values they bring, the skills they have, and the overall significance of this segment of the global society in general, the United Nations' Secretary-General appointed in 2017 an Envoy on Youth: Jayathma Wickramanayake from Sri Lanka. The Envoy on Youth's – who is naturally a young activist herself – has as her primary responsibilities:

> to work to ensure the participation of young people in issues that matter to them, giving them a voice at the United Nations and around the world. Championing the SDGs, she also brings the work of the UN closer to young people around the world.[11]

Alongside the Envoy on Youth, 17 other youth leaders and activists from all around the world are working to engage young people in the SDGs, advocate for their achievement, and contribute to supporting the United Nations' advocacy efforts to mobilize young people. The United Nations identifies five roles for the youth in the SDGs overall efforts as critical thinkers, change-makers, innovators, communicators, and leaders.[12] These and other similar measures taken by the UN and its different agencies are surely positive steps in the right direction.

We are also aware of several other international organizations that are becoming more aware of the important role of the youth and actively working to enhance youth participation in their interventions. Nevertheless, we observe a rather weak implementation of such initiatives on the country level. For example, we know that Voluntary National Reviews of multiple countries lack youth participation and voice. In the United Nations Synthesis Report on the VNRs after the 2018 High-Level Political Forum, for example, youth were mentioned by 46 countries that had conducted the VNRs in two main ways: a) as beneficiaries of SDGs-related policies

11 https://www.un.org/sustainabledevelopment/youth/
12 Ibidem

and programs; and b) as a group to focus on to support the implementation of the SDGs (Voluntary National Reviews 2018). There was no mention in the report and the countries' presentations specifically on the role of youth in the VNRs overall or on their countries' teams leading the SDGs implementation. Moreover, youth are not systematically involved in implementing and reviewing – including monitoring and evaluating – countries' progress in achieving the SDGs. This issue must be taken seriously by the United Nations and all its agencies when working with national SDGs teams especially through the VNR process.

EvalYouth is currently preparing for a deeper analysis of the VNRs from this perspective, to both document the extent of the problem and raise awareness regarding the roles that youth can and should play in these reports. Since its inception in 2015, the EvalYouth network has been working closely with young people and young evaluators specifically and with several other entities and agencies that are striving to enhance the role of young people. During the past few years, we observe prominent progress in this regard, but we also see that much more is needed to involve youth in truly meaningful ways. We observe, however, that the most significant force behind changing the status quo is young people themselves and hence the reference to revolution in this chapter.

Conclusion

In this chapter, we argued that a revolution – within and outside of evaluation – is taking place and that having youth at the table is more paramount than any time before in order to realize the SDGs. Evaluation *done well* needs to be revolutionary and, for this reason, it needs to include the powerful force of young people. The role of young people must be as collaborators and co-leaders, not as objects of evaluation. Consequently, youth participatory evaluation, which means moving from evaluation *on* youth to evaluation *with* and *by* youth, offers great opportunities to strengthen our global efforts portrayed in the Agenda 2030.

Involving young people in conducting evaluations not only offers a viewpoint from the perspective of one of the largest segments of the world's population, but it also brings many potential benefits. First, by and large, young evaluators bring social justice, equality, equity, peace, and environmental sustainability values. Second, young people bring cutting-edge technology and communications skills to evaluation practice. Lastly, young and emerging evaluators are a vital supply of workforce that is eager to contribute to developing the field.

Some necessary measures have been implemented by the United Nations and its agencies to seriously include young people in the SDGs. Yet, we observe a neglect at the country level in this regard. The United Nations with all its agencies must work closely with all countries' SDGs teams and other national institutions to further promote the role of young people in these processes.

The evaluation community must also decide what role they want to play. Does the evaluation community want to continue with the status quo by continuing to marginalize youth and young and emerging evaluators? Alternatively, will the evaluation community advocate and create space for youth and young and emerging evaluators as collaborators and co-leaders? Silence or apathy is not an option.

We reiterate our call to action. The revolution is coming. In many ways, it is already here. What role will history say you played in this revolution?

References

Camino, Linda. 2005. Pitfalls and promising practices of youth-adult partnerships. An evaluator's reflections. *Journal of Community Psychology* 33 (1): 75-85. https://doi.org/10.1002/jcop.20043

Checkoway, Barry, and Kathie Richards-Schuster. 2003. Youth participation in community evaluation research. *American Journal of Evaluation* 24 (1): 21-33. https://doi.org/10.1177/109821400302400103

Guy, Jack, Hellen Regan. 27 May 2018. 8 key takeaways from the European election 2019 results. *CNN*. https://tinyurl.com/y272yop8.

Gray, Jasmin. 8 August 2019. *A third of young voters now believe the army should run the country*. Huffpost. https://tinyurl.com/yxl3ll4v.

Cousins, J. Bradley. (In press). *Collaborative approaches to evaluation: applying and assessing evidence-based principles*. Thousand Oaks, CA: Sage (Evaluation in practice series).

Cram, Fiona, Katherine Tibbetts, and Joan LaFrance. 2018. Indigenous evaluation. In: *New Directions for Evaluation* 2018 (159): 1-133. https://tinyurl.com/y5q48oao.

Hood, Stafford, Rodney Hopson, and Henry Frierson. 2005. *The role of culture and cultural context in evaluation. A mandate for inclusion, the discovery of truth and understanding*. Greenwich, CT: Information Age Publishing.

Hulshof, Henk. 2019. Child participation in evaluating social protection projects. Do global development actors walk the talk? *Progress in Development Studies* 19 (1): 1-20. https://doi.org/10.1177/1464993418805170.

Lattimer, Charlotte, Sophia Swithern, Dan Sparks et al. 2018. *Global Humanitarian Assistance Report*. Bristol, UK: Development Initiatives. https://tinyurl.com/yysok2ga.

Laws, Sophie, and Gillian Mann. 2004. *So you want to involve children in research? A toolkit supporting children's meaningful and ethical participation in research relating to violence against children*. Stockholm: Save the Children. https://tinyurl.com/yxscuspa.

Lisboa+21 Declaration on Youth Policies and Programmes 2019. Adopted at the World Conference of Ministers Responsible for Youth 2019 and the Youth Forum Lisboa+21 in Lisbon, 23 June 2019. https://tinyurl.com/y4baunbs.

London, Johathan K., Kristen Zimmerman, and Nancy Erbstein. 2003. Youth-led research and evaluation. Tools for youth, organizational and community development. *New Directions for Evaluation* 2003 (98): 33-45. https://doi.org/10.1002/ev.83.

Mertens, Donna M. 1999. Inclusive evaluation: Implications of transformative theory for evaluation. *American Journal of Evaluation* 20 (1): 1-14. https://doi.org/10.1177/109821409902000102.

Mertens, Donna M., and Amy T. Wilson. 2018. *Program Evaluation Theory and Practice*, 2nd ed. New York: Guilford Press.

Patton, Michael Quinn. 1978. *Utilization-focused evaluation*. Beverly Hills, CA: Sage.

Patton, Michael Quinn. 2017. A reading guide. *In: Evaluation for Agenda 2030. Providing Evidence on Progress and Sustainability*, (eds) Rob D. van den Berg, Indran Naidoo, and Susan Tamondong, XI-XVIII. Exeter: IDEAS. https://ideas-global.org/evaluation-for-agenda-2030/.

Purdue, Sophie, Heidi Peterson, and Christine Deng. 2018. The case for greater youth participation in monitoring and evaluation in international development. *Evaluation Journal of Australasia* 18 (4): 206-221. https://doi.org/10.1177/1035719X18804401.

Richards-Schuster, Kathie, and Sara Plachta Elliott. 2019. A practice matrix for involving young people in evaluation. Possibilities and considerations. *American Journal of Evaluation*. https://doi.org/10.1177/1098214019832113.

Segone, Marco, and Jim Rugh (eds). 2013. *Evaluation and Civil Society. Stakeholders' perspectives on national evaluation capacity development. Evaluation Working Paper.* New York: UNICEF. https://tinyurl.com/y6e7ount.

Segone, Marco, and Florence Tateossian. 2017. No one left behind. A focus on gender and social equity. In: *Evaluation for Agenda 2030. Providing Evidence on Progress and Sustainability*, (eds) Rob D. van den Berg Rob D., Indran Naidoo, and Susan Tamondong, 23-24. Exeter: IDEAS. https://ideas-global.org/evaluation-for-agenda-2030/.

Spratt, Vicky. 2018. The truth about young people and Brexit. We reality check the whole shebang. *BBC* https://tinyurl.com/y8wzrtyo.

UNDP. 2017. *Youth as partners for the implementation of the SDGs. Fast Fact briefing note*. New York: UNDP. https://tinyurl.com/yyq64rw5.

UNDP. 2018. *Human development indices and indicators. 2018 Statistical update*. New York: UNDP. http://report.hdr.undp.org/.

UNICEF. 2018. *UNICEF guidance note. Adolescent participation in UNICEF Monitoring and Evaluation*. New York: UNICEF Evaluation Office. https://tinyurl.com/yxfm5m5x.

United Nations. 2015. *Paris agreement.* New York: United Nations. https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

United Nations General Assembly. 14 December 1995. *World programme of action for youth to the year 2000 and beyond.* A/RES/50/81. https://tinyurl.com/y4btgl5f.

United Nations General Assembly. 25 September 2015. *Transforming our world. The 2030 Agenda for Sustainable Development.* A/RES/70/1. https://tinyurl.com/z4o6pk4.

UN Women. 2019. *Families in a changing world. Progress of the world's women 2019-2020.* New York: UN Women. https://tinyurl.com/y4kvu3a3.

Voluntary National Reviews. Synthesis Report. 2018. Coord. by United Nations Division for Sustainable Development Goals and Department of Economic and Social Affairs. New York: United Nations. https://tinyurl.com/yxm43b6n.

World Economic Forum. 2018. *The global gender gap report.* World Economic Forum, Cologne/Geneva, Switzerland. http://www3.weforum.org/docs/WEF_GGGR_2018.pdf.

Wridt, Pamela. 2018. Young people's participation in program design research, monitoring, and evaluation. In: *Handbook of Adolescent Development Research and Its Impact on Global Policy*, (eds) Jennifer E. Lansford, and Prerna Banati, 432-453. Oxford: Oxford University Press.

Youth 2030. The United Nations strategy on youth. 2018. [New York: Office of the Secretary-General's Envoy on Youth]. https://tinyurl.com/y56876bu.

Zeldin, Shep, Libby Bestul, and Jane Powers. 2012. *Youth-adult partnerships in evaluation (y-ap/e). A resource guide for translating research into practice.* Ithaca, NY: ACT for Youth Center of Excellence at Cornell University. https://tinyurl.com/yy77fl7b

CHAPTER 4

Emerging issues in national monitoring and evaluation systems in Africa and in Latin America

Abdoulaye Gounou and Gabriela Pérez Yarahuán

ABSTRACT. The landscape of monitoring and evaluation (M&E) in Africa and Latin America has changed substantially in the past two decades. While there is no uniform progress in all cases, there is clear evidence that some countries have transformed government structures and national legislation to imbed M&E requirements. Other countries are still at an early stage of development, but overall there appears to be a growing interest in and demand for an effective and functional National Evaluation Systems (NES) and Policy (NEP). This chapter portrays some of the developments that have taken place and analyses how these transformations will be shaping the capacity of government to respond to national and international demands, especially in light of 2030 Agenda. The chapter is divided into four sections; the introduction describes in general terms the situation, needs and expectations in the conformation of NES; the second and third sections address the recent situation in Africa and Latin America. The last section contains a set of final remarks that address some of the challenges for M&E in the current regional and international contexts.

Introduction

During the last two decades governments and civil society organizations in the developing world have experienced important changes in the way policies, programs and interventions get designed, implemented, monitored and evaluated. The transformations on government aim at building monitoring and evaluation systems that aid decision-makers in implementing policy and program with better results.

The influence of the international community, international cooperation agencies and multilateral financial organizations, has been relevant to the changes that have taken place. However, the specific political, economic, societal and cultural contexts have affected how countries have adapted to the new demands for monitoring and evaluation requirements. A significant number of countries in the South have experienced waves of innovation and learning in this new paradigm of government action, but we find that these experiences are not documented periodically, losing valuable information and learning. In Africa and Latin America there are countries, such as Chile, Colombia, Mexico, South Africa, Uganda and Benin, that have advanced substantially in the construction of national M&E systems. Countries like Argentina, Costa Rica, Peru, Uruguay, Ghana, Niger, Kenya and Botswana exhibit relative progress, while others have found obstacles for change.

The implementation of National Evaluation Systems (NESs) generates diverse expectations among different stakeholders. Government officials anticipate inputs to make policy decisions, but also to report to national and international audiences, including the submissions of Voluntary National Reviews (VNRs) to support progress on the Sustainable Development Goals (SDGs). Citizens and the media expect to be informed to hold governments accountable. Donors and financial institutions are eager to see results from their funding strategies. However, fulfilling every expectation has proven to be no easy task.

There are rising demand and expectations to institutionalize NES in Africa and Latin America. However, it appears that in many cases evaluation results or results-based systems implementation may focused on reinforcing control and oversight rather than learning agendas for the various stakeholders involved in policies. The likely implication for the 2030 Agenda may be that the SDGs' review and follow-up process will be more data-driven than evaluation led. The risk for the 2030 agenda is that VNRs take the form of data reports on the great number of indicators included in the SDGs, gathered through the collection of data enforced

in government interventions, but fall short on promoting awareness on the needs of evaluation, evidence generation and learning.

There is no standard recipe that works for building NES. M&E institutions depend on governments' configurations, power structures and the nature of their political systems. In weak democracies and frail public administrations, M&E is a difficult endeavour. However, for the past decade an increasing interest, driven by economic crisis, the emergence of democratic systems and economic and social globalization, has resulted in growing exchanges among governments in the South. Learning has become more accepted among decision makers in charge of building a NES. It has resulted in some adaptation and innovation.

The critical transformation that the SDGs agenda demands from government and society requires that we understand the development, successes and failures, and what has been learned in the past two decades in the nature and configuration of NESs in the developing regions.

This chapter describes the experience in Africa and Latin America, two regions where much attention has been placed in the development of M&E capacities necessary for the strengthening of NESs, looking especially on how these experiences can be leveraged to deal with the requirements of Agenda 2030.

Africa

National evaluation policies and national evaluation systems

National evaluation policies are an important framework to structure, systematize and guide monitoring and evaluations at country level. Mwaijande (2018) cautions that the absence of such policies can potentially leave programs and policy planning unchecked and allow for inefficiency and ineffectiveness in policies and development program implementation and accountability.

It is upon this basis that this reflection advocates for the formulation and implementation of national evaluation policies across the continent. Højlund defines a national evaluation policy as a systematic and institutionalized monitoring and evaluation framework "in several interdependent organizational entities with the purpose of informing decision-making and securing oversight" (Højlund 2015, 36).

Mwaijande (2018) also defines a NEP as that which guides the evaluation process, activities, resources, and utilization of evaluation results. There is a direct link between national evaluation policies and national evaluation systems. A NES is put in place to implement a NEP; they complement each other. NEPs provide a normative framework, while national evaluation systems build the mechanisms that operationalize the principles dictated in the policy. However, the presence or absence of evaluation policies or systems does not necessarily hinder the undertaking of evaluations on the continent, as seen in Burkina Faso, Côte d'Ivoire, Niger, and Senegal. In many countries, evaluations are conducted by governments with the support of donors or are carried out by donors for their own accountability and learning purposes. The existence of more and more evaluation policies and systems help to better structure and understand the way evaluations take place. This is particularly demonstrated in South Africa and Uganda where government departments carried out evaluation before a National Evaluation Policy framework was passed. But sometimes, joint evaluations are carried out by governments and donors. In Benin a joint evaluation was implemented in 2014 with the World Bank.

While NEPs and NESs are supposed to increase the demand for and use of evaluations, governments do not necessarily need both to be in place to carry out evaluations. This is a key lesson for countries establishing their evaluation systems (Goldman et al. 2018), drawn from experiences in Benin, South Africa and Uganda. These countries are seeking to build a wide evaluation ecosystem, essential for institutionalization of evaluation that works with national evaluation associations, academia, civil society organizations, etc. They are seeking to put in place a system that allows evaluation to furnish evidence to policy-makers, notably around national capacity building in evaluation.

Twende Mbele cooperation and achievements

All three countries cooperating in the Twende Mbele programme[1] have a NEP. In South Africa, it was developed before the NES; in Benin and Uganda both countries were in the process of implementing a national system before "developing a policy". All three have developed an evaluation agenda or plan to prioritize evaluations for each year. Key components of policy and system are assembled in Table 1.

1 Twende Mbele: a three years peer learning cooperation programme between South Africa, Benin and Uganda with the financial support from DFID; WACIE: West Africa Capacity building in Impact Evaluation with the support of 3ie.

TABLE 1. Comparison of the three Twende Mbele countries

COMPONENTS	Benin	Uganda	South Africa
Eval Policy	Yes	Yes	Yes
Plan of what to evaluate	Strategic Evaluation Plan 2013–2015	3-year rolling evaluation agenda in place, indicating sector, topic and why the evaluations	Annual national, 7/9 provincial and emerging 44 departmental plans
Type of evaluations	14 implementations and/or process evaluations, 1 impact	Implementation, 4 process evaluations and 3 impact evaluations	45 implementations (process, some summative), 8 impact, 5 diagnostic, 1 economic
Institutional structure for coordination and oversight	BEPPAG, Presidency. National Evaluation Board selects evaluations, and involves range of stakeholders	GEF in OPM National Monitoring and Evaluation Technical Working Group	DPME overall 'owner' of system. Supported by Evaluation Technical Working Group involving national and provincial departments
Coordination with donor M&E mechanisms	There is a platform for this in the Ministry of Planning	M&E Department relates with the donor economists group, and the Donor Partnership Forum	Donors do not play a big role and so focus is on the government
Nos. where high degree of implementation (%)	6/9 (67%). In follow-up observe significant changes from implementation of findings	Follow-up shows between 10% and 30% of evaluation recommendations have been taken up	9/16 have implemented > 25% of recommendations; in 2 the improvement plans have been incorporated into another evaluation

SOURCE: COMPILED FROM GOVERNMENT DATA IN GOLDMAN ET AL. 2018.

The evaluation system is separated from regular monitoring and data gathering functions in all three countries. Mechanisms for promoting autonomy and impartiality of evaluation have been developed, including the important role of the central evaluation unit in managing the interface between supply (undertaking quality evaluations) and demand from central policy units.

All three countries use independent service providers for reasons of independence and/or impartiality, as well as lack of capacity in government to actually undertake evaluations. They have a system for dissemination, but this is still relatively technocratic and can be enhanced to increase knowledge of evaluation results,

in government, Parliament and the public. The three Twende Mbele countries are examples for other countries in Africa to be inspired by.

Evaluation units, located within public institutions, are important actors responsible for the production and dissemination of evaluative knowledge in complex programming and institutional settings. In all three countries M&E units exist in both national departments and municipalities.

All three countries have a National Development Plan (NDP), indicators which are monitored, and which are supposed to link to departmental/ministry plans and that departments have indicators and targets which reflect progress integrated into their annual plans (Goldman et al. 2018). They undertake routine monitoring of performance. In the case of South Africa, quarterly reporting on annual performance plan is due to the Cabinet, and in Uganda the same exercise is done on key priorities and the information needed for management and planning.

The institutionalization has been facilitated in the countries by high level political will among the government, the high level location of the Unit in charge of M&E and policy promotion (Presidency in South Africa and Benin and the Office of the Prime Minister in Uganda), the existence of the NEP (South Africa and Benin) and M&E Strategy in Uganda, strong leadership at country level, institutionalization of evaluation associations and meetings (SAMEA in South Africa, JBE in Benin and UEW in Uganda) and dynamic partnerships in evaluation, as witnessed in Twende Mbele. This increased the national awareness and ownership. Evaluations are better conceptualized financed, conducted and completed more easily with more professionalism and supporting evidence-based government.

In Uganda, a National Integrated Monitoring and Evaluation Strategy (NIMES) has been introduced in 2005 sought to strengthen performance assessment in the public sector. The national M&E policy of 2013 updated the NIMES of 2006, with a government evaluation agenda or plan including eight evaluations. An implementation plan was developed in 2013/2014, along with the first evaluations through the facility. The Government Evaluation Facility (GEF) was established in the Office of the Prime Minister (OPM) in 2013. It received two mandates: a) design, conduct, and commission and disseminate evaluations on public policies and major public investments and b) oversee improvements in the quality and utility of evaluations conducted across government at a decentralised level.

In Uganda, a total of 65 national evaluations have been completed or are underway, with provincial and departmental evaluations. In terms of use, the OPM

required to provide 6-monthly briefings to the Cabinet or a designated Cabinet Sub-Committee on the status of evaluations underway and findings of evaluations when they are complete. A government response and implementation tracking mechanism developed by the OPM has been put in place to establish how many recommendations of the evaluation findings have been implemented. So far, a follow-up by the OPM has found that about 30% of recommendations are being implemented. This demonstrates that the promotion of evaluation culture and the use of evaluation findings are still a great challenge for African countries. The first provincial evaluation plans were piloted in 2012/2013 and the system has gradually widened now to include departmental evaluation plans.

In South Africa, a Ministry and Department of Performance and M&E was established in 2009, which changed to the Department of Planning and M&E (DPME) in 2010. A National Evaluation Policy Framework (NEPF) was approved by the Cabinet in 2011 and in the same year, an Evaluation and Research Unit (ERU) was established in the DPME to develop and run the evaluation system.

Experiences in Benin

In the case of Benin, evaluations findings are seen as "public goods" available through a website.[2] There is still a challenge to build the links between evaluation and planning and budgeting. This is due to bureaucratic issues, evaluation culture at country level, and capacity and ownership issues despite the existence of a strong political will in Benin. In 2018, the National Methodology Guide for Evaluation introduced Theories of Change in all public initiatives that require financial support from the national Budget.

The institutional design of Government M&E systems is important, including systems for capturing, processing, storing and communicating M&E information. Monitoring helps managers and policymakers understand what the money invested is producing and whether plans are being followed. Evaluation helps to establish what difference is being made, why the level of performance is being achieved, what is being learned from activities, and whether and how to strengthen the implementation of a programme or policy.

In 2008 a Bureau of Public Policies Evaluation was established in the Ministry

2 See www.presidence.bj/evaluation-politiques-publiques.

of Planning of Benin. This Bureau was hosted by the General Secretariat of the Presidency under the name of the Bureau of Public Policies Evaluation and Government Action Analysis. This was due to the establishment of a competencies framework for different actors related to evaluation under the responsibility of the presidency of the Republic of Benin and in view of promoting good governance.

The role of the Benin Bureau of Public Policy Evaluation is to establish and lead the National Evaluation System (NES), ensure evaluation becomes a strategic management tool for development and commission evaluations whether demanded by donors, national government or by the local government. A national evaluation policy was adopted in 2012 and an institutional framework established defining the mechanisms for conducting evaluations including guidance on selecting evaluations and structures, engagement of stakeholders, dissemination of results and the monitoring of implementation of recommendations. To assist with impartiality, independent service providers undertake the evaluations, such as university consultancy firms as well as independent consultants. The Bureau of Public Policy Evaluation commissioned and completed evaluations including sectoral projects, multisectoral programmes and public policies in decentralisation, power, agriculture, health, water and energy and specific studies in the areas of domestic electric security policy and global development strategy.

In terms of the use of evaluations, a study was undertaken by the Bureau of Public Policies Evaluation on quality and use of evaluations commissioned from 2010 to 2014 in Benin, focusing on nine evaluations. One of the key findings was good ownership of the recommendations by implementing agencies. Approximately 80% of the recommendations (from all nine evaluations) have led to the development of implementation plans. Approximately 82% of the recommendations led to specific changes (49% policy review, 10% institutional change, 10% new projects and 15 other short-term measures). However, it is an ongoing challenge to ensure the use of evaluation findings for policy improvement and better implementation.[3]

Challenges for the future

At the country-level, we see an increasing collaboration between government, civil society and academia on evaluation. However, different streams of support, and

3 Direction Générale de l'Evaluation (Bureau of Public Policy Evaluation) : Rapport sur le suivi de l'utilisation des recommandations de l'évaluation au Bénin, edition 2015.

different levels of engagement – at times duplicative – around eitherare present. There are many opportunities to collaborate more closely to strengthen capacities and evaluation culture at the national level which means the involvement of academia (curricula development and research), VOPES (conducting evaluation) and civil society (for dissemination).

For further efforts to strengthen National Evaluation Systems, countries need legitimacy from the parliament through legislation. In Benin, the Government is willing to examine whether to include evaluation in the Constitution. Participants of the August 2018 Benin Evaluation Days held in Cotonou, recommended that evaluation must be integrated into the Constitution. Furthermore, a law on evaluation is being designed in case the proposed modification of the Constitution fails.

The demand for National Evaluation Policies (NEPs) has gained traction in Africa as more countries acknowledge the value of evaluation in enhancing efficiency, effectiveness, and equity in public and development management. As these countries implement developmental programmes and projects, they require appropriate and effective monitoring and evaluation systems to measure performance, assess impact and draw lessons for future programmes and projects. Similarly, the Sustainable Development Goals call for country-led systems to measure effectiveness, efficiency, relevance, sustainability, and the impact of development interventions (Chirau, Waller and Blaser-Mapitsa 2018).

African perspectives on Voluntary National Reviews (VNR)

Countries are beginning to document their progress towards the Sustainable Development Goals (SDGs) through regular Voluntary National Reviews submitted to the UN's High-Level Political Forum on Sustainable Development. The 2030 Agenda for Sustainable Development sets clear expectations for VNRs to be "rigorous and based on evidence", and the UN recommends that the first VNR in particular should describe the reviewing process and how the "national follow-up and reporting" system will be implemented. Evaluation enhances monitoring's meaning and depth by addressing complexity in how the SDGs are best achieved so each VNR should include up-to-date evaluation findings and an assessment of progress on national evaluation policies and systems.

At the 2016 UN High-Level Political Forum, 22 countries presented Voluntary National Reviews (VNRs) — status reports on their efforts to implement national-

level follow-up and review frameworks for the Sustainable Development Goals (SDGs) – among which, Benin. The analysis of the 22 VNRs, which focused on how each addressed the role of evaluation, indicates that most VNRs show little awareness about what evaluation is and how it could be used to support the 2030 Agenda and even more so regarding the agenda 2063 from African Union.

In most African countries, monitoring is strong, but evaluation systems and processes often remain missing or misunderstood for their role in the SDGs. Good practices are emerging across the world, such as: linked National Evaluation Policy and action planning (Nepal); recognition of the SDGs' complexity when considering evaluation (Czech Republic); learning through evaluation to feed VNRs report (Ethiopia and Kenya); and drawing on findings from past evaluations (Belize). Countries still to submit their first VNR could build on these examples.

Countries like Ghana, Kenya and Niger are putting in place their National Evaluation System. The lack of qualitative data limits the interpretation of these results. It is critical is to see the trends that are emerging and to understand how some of these findings are understood and interpreted by the respondents. For all this, qualitative information is crucial.

The pressure to deliver results often limits the amount of time available to define the root cause of a problem. This is a false economy, as it can lead to projects and programmes that address symptoms rather than causes. The current VNR processes across the continent represent a unique entry point at national level thinking on evidence-creation and dissemination for transformative changes. Benin should seek to embed evaluation into national evaluation plans and policies to foster transformative development, but monitoring data seems to prevail within the scope of the VNR processes. Evaluation is often dismissed.

For transformative change to take place, the generation and use of evidence are critical. However, such evidence ought to be as diverse as possible. The Sustainable Development Goals are not about quantitative data only but also qualitative. Measures that could greatly strengthen the effectiveness of VNR processes include the following: the conduct of meta-analyses and joint evaluations, the launch of evaluations of cross-cutting strategies and the support of VOPE, CSOs and parliaments.

When it comes to the SDGs, no government can do it alone. Partnerships across a wide range of national and international players are essential to make an effort to set up a platform where all the interested parties could collectively bring the discourse on VNR further, both at national and sub-national levels.

Benin has a robust National Evaluation System so its VNR should be strongly supported by the evaluation findings. Unfortunately, this is not actually the case. Statistical data and administrative information are prioritised. In addition, the country has different actors involved in the VNR process who are not (yet) really collaborating.

In Africa, M&E is seen by many as a role of the M&E Unit rather than of all managers – a viewpoint expressed by 58.8% of respondents in South Africa, 63.8% in Benin and 54.7% of managers in Uganda, suggesting that monitoring is not yet embedded as part of managers' roles. However, a positive sign is that those responsible for M&E are positioned at a high level, with around half respondents saying that the most senior people responsible for M&E are Deputy Directors-General or their equivalents, with the other half indicating that these were at Director level or below. This indicates:

a. A strengthening of the role of evaluation, and the increasing integration of an evaluation perspective in planning, monitoring and reporting (especially country-level reporting, as well as for the Agendas 2030 and 2063);

b. The need for governments, in collaboration with their partners, to plan for evaluation, and set aside funding for M & E. This requires a change in budgetary culture and practices, as well as political will, and

c. The need for strengthening robust quantitative data to monitor developmental processes and include all stakeholders in this process (beyond default household surveys conducted by national statistical offices).

Latin America

The impetus for the expansion and institutionalization of national evaluation systems (NES) in Latin America came in the context of the repeated economic crisis and democratic transitions particularly in the last two decades of the 20th Century. The architecture of most of the national evaluation systems in the region has privileged control and accountability in detriment of dialogue, participation and learning. The challenge today is how to take advantage of the progress made and at the same time imbed an evaluation culture that helps to implement more effective policies to deal with the complex and interrelated objectives of the 21st Century highlighted by the 2030 international development agenda and the fulfilment of the Sustainable Development Goals (SDGs).

The first stage of government monitoring and evaluation of public programs in Latin America, back in the 1970s, was mostly driven by international financed projects in the rural sector, particularly in Mexico and Brazil that served as a demonstration effect of evaluation´s potential in other countries (Feinstein 2012). By the late 1980s recurrent economic crisis in Latin American countries had made evident that decisions on policy mattered more than ever before as the constraints on resources became palpable and results on important goals had not been obtained. The problems of rising poverty and inequality became manifest (Gasparini and Cruces 2013) and consequently the pressure on governments to deliver on promises made. On the political front, democratic transitions became a reality, after years of popular pressures and citizen demands, at least in terms of free elected governments and political party competition. People´s movements and civil society organizations started to take shape to express more structured policy demands (Donaghy 2018).

The changes in the political and economic circumstances required governments to adjust to a new public administration paradigm, to include performance measurement, accountability and transparency policies and institutions. In this context, several countries began to build the institutional underpinnings of national evaluation systems, some as early as 1991 (Colombia). Laws were enacted, regulations were set in place, and in some cases evaluation functions were located in specific governmental entities, though with different approaches. In Colombia evaluation was placed at the Ministry of Planning (DPN); in Chile at the Finance Ministry (DIPRES), as well as in Peru, at the National General Directorate of Public Budget (DGNP), both countries also have as important stakeholders and allies the Ministries or Departments of Social Policy. In Mexico, at first functions were fragmented in several government agencies (with the predominance of the Ministry of Social Development) and since 2004 an evaluation council for social policy (CONEVAL) was created. In Argentina and Brazil, the evaluation function was placed at the Office of the Presidency, in the latter alongside the Ministry of Social Development (SAGI).

Today most governments in Latin America formally recognize monitoring and evaluation activities as a requirement in the policy-budget cycle (Pérez Yarahuán and Maldonado Trujillo 2015). And while monitoring of government programs is recognized as a regular activity carried out by public agencies, evaluation is generally conceived as an effort undertaken by experts outside the government sphere to guarantee independence and shield it from conflicts of interest.

In practice, progress has not been homogeneous in the region, either across countries or through time. Evaluation systems can be described by their advancement in a set of components, among the most relevant are:

1. Awareness and recognition of the evaluation function (awareness by stakeholders and establishment of formal mandates);

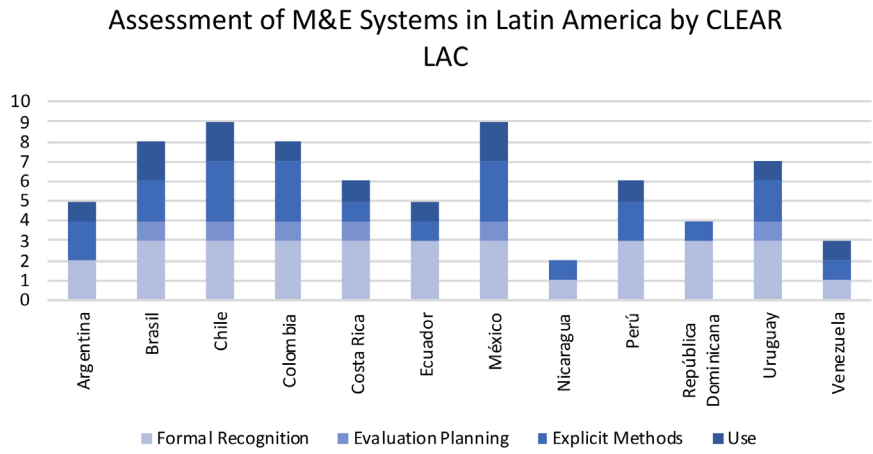
2. Evaluation process planning (transparent communication of what is to be evaluated and when);

3. Transparency on methods and reliable information used in evaluations; and

4. Utilization of evaluation results to develop effective policies.

An analysis of these components was made in 2015 (Pérez Yarahuán and Maldonado Trujillo 2015) taking into account ten country case studies. The following figure shows a representation on the situation of National M&E systems by 2015.

FIGURE 1. Assessment of M&E systems in Latin America



SOURCE: PÉREZ YARAHUÁN, AND TRUJILLO 2015, 404.

As stated previously, awareness and formal recognition of the fundamental role of evaluation as part of the policy cycle is high in most Latin American countries. However, recent political circumstances might hinder or put at risk the prominence held until now of explicit M&E policy. In Mexico, one of the countries with the greatest progress in this area, it appears that the administration, led by President Andrés Manuel López Obrador since 2018, has put little or no attention to previous evaluation results in the design of its social policy agenda. To illustrate this, take the most evaluated program, the conditional cash transfer program (PROSPERA) which made international headlines and was the inspiration of several similar programs across the world. This program was recently cancelled without any regards to its

numerous evaluations that stated positive results in key objectives.[4] This decision by the current Mexican government may be a signal of a backlash effect, a result of the shortcomings in the implementation of NES. This unsettling decision on what seemed an upward trend in the institutionalization of evaluation, may be the opportunity to rethink the design and architecture of M&E systems in the region.

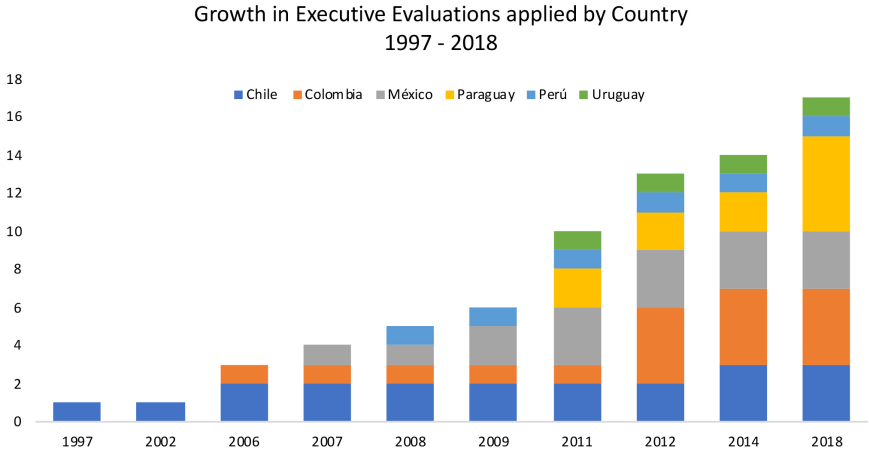
In some countries, an evaluation cycle of government policies has been well established. This includes the process of decision making of what, how and when different governments' actions, interventions or programs will be evaluated. The importance of this cannot be underestimated, as it reduces uncertainty in the evaluation process, connects the government policy agenda and planning with evaluation requirements and it contributes to transparency. This has clearly been the case for Chile, Mexico and Colombia. However, for many countries in Latin America, the decisions of what programs get evaluated, when and how, does not follow a set of an established and clear set of rules.

With respect to evaluation methodologies and frameworks, Latin American countries for the most part have adopted, on the one hand, a traditional approach, including evaluation methods to generate expert external assessments for the design, process and impact of interventions and programs. On the other hand, several governments have implemented a hybrid type of evaluation, that delivers information on several aspects of a program (including its design, process and results) and it is done in a relatively short amount of time. These executive or rapid evaluations have been applied at least in Chile, Colombia, Costa Rica, Mexico, Paraguay, Perú and Uruguay. The advantages of this type of evaluation are that it delivers useful monitoring information, as well as an informed and expert external assessment on the programs evaluated. Some of the drawbacks of this type of evaluation are that they involve limited contact with stakeholders, such as program officials, operators and decision makers, and almost none with program beneficiaries. Evaluations of government programs that include a participatory approach are still quite scarce in Latin America's NESs. There are two possible explanations for this. First, at the beginning there was a need to guarantee specific standards and quality in evaluations and thus fixed traditional methodologies and strict terms of reference were put in place. Second, national evaluation capacities were deemed insufficient to cope with the increasing demand which made national governments uneasy of evaluations that entailed a more flexible approach and more time to perform. The following figure shows how many countries have

4 See https://tinyurl.com/yx9z56p6 (in Spanish).

adopted this type of evaluation in recent research conducted by Robert Kaufman and Mauricio García Moreno (forthcoming).

FIGURE 2. Growth in Executive Evaluations by country 1997-2018



SOURCE: FROM GARCÍA MORENO, MAURICIO Y JORGE KAUFMANN, EVALUACIONES RÁPIDAS PARA MEJORAR EL DESEMPEÑO DE LAS INTERVENCIONES PÚBLICAS EN AMÉRICA LATINA Y EL CARIBE (FORTHCOMING).

The use of evidence to increase the effectiveness of programs and policies is, as in many other regions, the Achilles heel of evaluation. In Latin America the focus of evaluation use has been aimed largely at the budgetary process and much less on management, policy planning and even less on learning and integration. Chile and Mexico have set in place rules that ensure a follow up mechanism of evaluation results and recommendations. However, these mechanisms run the risk of getting distorted as the recommendations from external evaluators are taken as key inputs for audit processes and those accountable are seldom in a position to make meaningful changes in programs and policies.

In the past twenty years, the outlook of M&E in the region has changed dramatically. Capacities have been built, information on public programs has been gathered and systematized, program logic frames have been formulated, evaluations have been delivered. These changes have happened with various degrees of consistency and quality, across countries, sectors and time. But progress in the NES in the region for the most part is an undeniable fact. Notwithstanding the headway made, the weakest aspect of the M&E systems appears to be in the use of the evidence to promote fruitful dialogue among stakeholders, to advance in substantive changes in policy and to connect policy sectors to cope with ever growing complex social

problems. NES in Latin America must experience a transformation, a next step in development where more attention is placed in generating dialogue among stakeholders, allowing for pertinent evaluation approaches, establishing the difference between evaluation and auditing.

The challenges that the SDGs pose for developing countries require the adequate functioning of national statistical institutions to monitor progress on a wide variety and extensive number of indicators, and that NES operate as open learning environments in order to create synergies along interrelated policy sectors.

Final Remarks

For Africa and Latin America, monitoring and evaluation for programs and policies in the public sector is a relatively recent undertaking. In Latin America, evaluation started in localized sectorial policies, particularly in agricultural projects. In this region, only in the late 1990s M&E systems and policies began to take shape on a broader scale, encompassing a wide array of policies and programs, mostly at the national level. In Africa, efforts to build and implement evaluation systems and policies have taken place mostly in the first decade of the 21st century.

Since the 2010s, Benin, Uganda and South Africa have undertaken a significant effort to mainstream evaluations in the work of government, in a great diversity of political situations and with different resource constraints. Systems are emerging with a wide variety of components – policies, plans, standards, governance structures, etc., which involve a wide range of stakeholders in the evaluation ecosystem. These have to reflect local realities and challenges. There is considerable local innovation in how to establish these systems, and adaptive management as these systems develop – an example of 'Made in Africa' rather than mimicry of the Global North.

In terms of use, there is evidence of a significant portion of evaluations having recommendations implemented and we are beginning to see examples of integration with the budget process. We see an emerging process of innovation and piloting, building capacity and with an ongoing need for political will to ensure use of evaluation findings. The peer learning approach has already enhanced these systems, and the resources being made available through the Twende Mbele programme provide an opportunity to deepen this and to expand evaluation to other countries in Africa (Goldman et al. 2018).

The existence of an M&E Policy/Strategy as well as a legal/regulatory frame contribute to the institutionalization of the National Evaluation Systems and help to improve M&E culture development such as evidence-based management and evidence-based policymaking. On the other hand, the development of a strong and dynamic partnership in evaluation promotes M&E culture and good governance as demonstrated by Twende Mbele (Partnership in evaluation between South Africa, Benin, Ghana and Uganda). The same will probably happen with WACIE (Partnership in evaluation between 3ie[5] and Government of Benin for capacity building in impact evaluation within height west Africa francophone countries.

The experience in Latin America spans over 20 years. Throughout this time there has been important progress on embedding monitoring and evaluation into government processes. We find that M&E is widely recognized today as an important component of the policy cycle, by stakeholders in government but also among active civil society organizations and the media. Recognition is also at a formal level, with legislation being enacted to formalize and assure compliance with M&E requirements. There has been experience sharing between country governments, and evaluation associations that are of more recent creation. This has led to some common characteristics and good practices, but also to innovation that responds to specific political and administrative and cultural contexts. Such is the case of the so-called 'executive evaluation' that focuses on a rapid assessment of programs, particularly intended for budget decision-making, but which differs in time and methods used.

More recently, and as the political landscape in Latin America has changed, NESs may be facing obstacles to maintain its ascending trend. This drawback may have been caused in part because the development of NESs has mainly focused on accountability and budget decisions and has not been directed or designed for policy planning and its learning potential. This may prove to be the biggest challenge to fulfil the 2030 agenda. The necessary transformation in policy and programs that needs to take place to comply with the SDGs requires that NESs not only report data, on the numerous indicators agreed upon by the international community, but also informs on the many interactions and complexity of policy and programs. More in-depth evaluation capacity building needs to take place amongst decision-

5 3ie (International Initiative for Impact Evaluation) is in partnership in evaluation with the Government of Benin to develop capacity building in impact evalaution in West Africa francophone countries with the financial support of West Africa Economy and Monetary Union (WAEMU) and West African development Bank (BOAD).

makers and the development and implementation of participatory methods to include stakeholders outside the government sphere needs to materialize for teh NESs to become evidence instruments that truly help developing countries play an active and positive role in future global development.

References

Chirau, Takunda, Cara Waller, and Caitlom Blaser-Mapitsa. 2018. *The National Evaluation Policy landscape in Africa. A Comparison.* Johannesburg: Clear-AA and Twende Mbele. Policy Brief. https://tinyurl.com/y3wo9ty3.

Donaghy, M. 2018. Reforming the relationship between the state and civil society in Latin America. *Latin American Research Review* 53 (2): 388–393. http://doi.org/10.25222/larr.620.

Feinstein, Osvaldo. 2012. La institucionalización de la evaluación de políticas públicas en América Latina. *Presupuesto y Gasto Público* 68: 41–52.

Gasparini, Leonardo, and Guillermo Cruces. 2013. Poverty and inequality in Latin America. A story of two decades. *Journal of International Affairs* 66 (2): 51–63. http://www.jstor.org/stable/24388285.

Goldman, Ian, Albert Byamugisha, Abdoulaye Gounou, et al. 2018. The emergence of government evaluation systems in Africa. The case of Benin, Uganda and South Africa. *African Evaluation Journal* 6 (1): 1-11 https://doi.org/10.4102/aej.v6i1.253.

Højlund, Steven. (2015). Evaluation in the European Commission. For accountability or learning? *European Journal of Risk Regulation* 6 (1): 35-46.

Mwaijambe, Francis. (2018). Why should countries have national evaluation policies? In: *eVALUation matters*, Building supply and demand for evaluation in Africa. Volume 1, 18-29.

Pérez Yarahuán, Gabriela, and Claudio Maldonado Trujillo (eds). 2015. *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina*. Mexico D.F.: Centro de Investigación y Docencia Económias, CLEAR. https://tinyurl.com/y2j7ljs5.

CHAPTER 5

The transformative agenda for evaluation in Small Island Developing States: the Caribbean and the Pacific

LEAD AUTHORS Lennise Baptiste and Viliamu Iese

CARIBBEAN AUTHORS Valerie Gordon, Arlene Bailey, Nadini Persaud and Claudia Nicholson

PACIFIC AUTHORS Annika Rose Dean, Morgan Wairiu, Otto Navunicagi, Jope Tarai and Chris Peter Teva

ABSTRACT. Small island developing states in the Caribbean and the Pacific have a history of collaborative efforts to address economic development. The vulnerability of the two regions has increased in the wake of climate change patterns which have illuminated development gaps in the areas of food security, housing, health, education and security for their populations. The targets of the sustainable development goals indicate pathways that can be followed to transform the regions. The authors discuss the issues which diminish the utility and quality of evaluation work such as weak M&E systems, leadership of evaluations, the perceived low technical capacity of regional personnel to conduct evaluations, access to different types of stakeholders, project versus impact and outcome evaluation, and the financing of M&E activities. These authors proffer some solutions such as building an evaluation culture, garnering political will at the highest levels, defining evaluators' role and desired competencies, and sharing lessons learned to describe the role evaluation can play in the transformation of the regions and raise a few challenges posed by these in the short and medium term.

Introduction

The Pacific and the Caribbean regions represent two groups of countries, consisting mainly of small island developing states (SIDS), whose leaders have committed to the vision of the UN 2030 Agenda and the realisation of the Sustainable Development Goals (SDGs). Though the two regions are on opposite sides of the world, there is value in the transference of knowledge and best practices among countries in the areas of economic diversification and resilience to natural hazards (Sirtaine and Melanson 2018). The regions' vulnerability to climate change necessitates the diversion of significant percentages of GDP to facilitate post-disaster activities (Thomas and Benjamin 2018). Thus, there is an urgent need for transformative public sector thinking concerning monitoring and evaluation (M&E), and by extension, improving the utility and the use by the regions of evaluation findings from climate change and other sustainable development initiatives. In this chapter, the authors review current regional evaluation practices and discuss how evaluation can be leveraged to achieve the transformation intended by the vision of the 2030 Agenda to improve the quality of the lives of their citizens.

Early collaborations between the Caribbean and the Pacific regions have emphasised the importance of the integration of M&E mechanisms in the climate change adaptation initiatives in both regions (CARICOM 2004). The Caribbean Community Climate Change Centre (CCCCC) has been collaborating with the Secretariat of the Pacific Regional Environment Programme (SPREP) to encourage increased sharing of knowledge and experiences across the two regions. A Regional Clearinghouse Database coordinated by the CCCCC provides access to information on climate change projects in the Caribbean and the Pacific regions. The Commonwealth Secretariat has provided support to the CCCCC in the development of an M&E Framework, as well as support to enhance collaboration with the Pacific region (Commonwealth Secretariat 2019). The collaboration between the regions has continued with the coordination of workshops and implementation of projects such as IMPACT, which seeks to enhance the analysis of climate change data and facilitate comparability and transferability of findings across regions. Collaborations between Caribbean M&E experts and the University of the South Pacific's (USP) Pacific Centre for Environment and Sustainable Development (PaCE-SD) researchers led to the compilation of this chapter.

The goals and targets of the vision of the 2030 Agenda can be grouped according to identified areas of critical importance: people, planet, prosperity, peace, and partnerships. The priority given to any of these identified areas will depend on the

contextual issues in each country, and there are differences among the countries within these regions. These SIDS have been selective about the SDG indicators they will measure and the targets they will pursue. Though there was a commitment to achieving the vision of the 2030 Agenda, there was no documented plan for the systematic implementation of the practices to achieve the SDG targets and outcomes. The role of evaluation is therefore critical to document the progress made towards the intended transformation of the Pacific and the Caribbean regions.

Caribbean region

Geographically, the Caribbean comprises twenty-two island states in the Caribbean Sea from Cuba, Cayman Islands and Jamaica in the northwest, to Barbados in the east and the twin-island state of Trinidad and Tobago in the south. The islands of Bermuda, The Commonwealth to the Bahamas and Turks and Caicos in the Northern Atlantic, as well as Belize in Central America, Guyana and Suriname in South America are also considered part of the Caribbean. Additionally, located in the archipelago of islands are the independent countries of the Dominican Republic and Haiti, Puerto Rico which is an overseas territory of the USA, and the Virgin Islands which are shared between the UK and the US as overseas territories. Therefore, any consideration of the nature and state of evaluation in the Caribbean must acknowledge the diversity among the group of countries in terms of language, economies, and development. This variety results in the main from the administrative and economic relationships held with England, France, the Kingdom of the Netherlands, Portugal and Spain birthed from the region's colonial history. The Caribbean should therefore not be painted with one broad brush because there are unique cultural differences among the states. The population of this region comprises approximately 45.5 million people.[1] While the Caribbean region is located North of the equator and has closer proximity to the USA and Canada than the European countries previously mentioned, the Caribbean tends to be associated more with the Global South, a term not associated with location but with the development status of countries. The term Global South immediately references colonial history, neo-imperialism, differing economic and social advantage which maintain inequalities when living standards, life expectancy, and access to resources are considered (Dados and Connell 2012).

1 See database on: https://www.worldometers.info/world-population/caribbean-population/

The Caribbean is comprised of independent states and some overseas territories of the UK, Netherlands and the USA. Various economic partnerships among Caribbean countries have also resulted in the formulation of several groupings. The Caribbean Community (CARICOM) is focused on the four areas of economic integration, foreign policy coordination, human and social development and security. The Organisation of Eastern Caribbean States (OECS) is an international inter-governmental Organisation dedicated to economic harmonisation and integration, protection of human and legal rights, and the encouragement of good governance among independent and non-independent countries in the Eastern Caribbean. The English-speaking islands and the mainland nations of Belize and Guyana receive technical assistance from the Commonwealth Secretariat as they once constituted the Caribbean portion of the British Empire. The Forum of the Caribbean Group of African, Caribbean and Pacific (ACP) States (CARIFORUM) is focused on promoting and coordinating policy dialogue, cooperation and regional integration, mainly within the framework of the Cotonou Agreement between the ACP and the European Union and also the CARIFORUM-European Community Economic Partnership Agreement (EPA).

While historically regional development in the Caribbean was financed with colonial resources, in the last four decades funding from donor agencies has in the main supported the development of Caribbean countries in the areas of education, health, housing, agriculture, citizen security, and the environment. Donor agencies and mechanisms such as the United Nations Development Programme (UNDP), the United States President's Emergency Plan for AIDS Relief (PEPFAR), the United States Agency for International Development (USAID), the United Nations Children's Fund (UNICEF), the Pan American Health Organization (PAHO), the Inter-American Development Bank (IDB) are all represented in the region. The countries in this geographical location share several common challenges, primarily the need for increased global competitiveness and economic growth. Most of the countries in this region are also classified as highly indebted (Robinson 2014).

Pacific region

The Pacific Island Countries (PICs) are also heterogeneous and divided into three sub-regions including Melanesia (mostly in the central-western Pacific), Polynesia (central-eastern Pacific) and Micronesia (mostly low-lying islands in the northern Pacific). The region is made up of about 20,000 islands and atolls and is home to about 11 million people. This region is scattered across the Pacific Ocean, the

largest ocean covering one third of the Earth's surface. The central importance of the ocean to Pacific Islanders led the late Professor Hau'ofa to refer to the Pacific as "a sea of islands", describing the ocean not as a barrier between countries, but as a source of connection, sustenance and cultural identity – "we are the sea, we are the ocean" (Hau'ofa 1993). The larger Melanesian countries of Papua New Guinea, Fiji, New Caledonia, Solomon Islands and Vanuatu represent 90% of the total Pacific Islands' population and 85% of the total land area. The region also includes some of the smallest and lowest elevation countries in the world. For example, the landmass of Tuvalu is only 26km^2, with its highest point being just 5m above sea level (Holland et al. 2018; Wairiu et al. 2012). Heterogeneity exists both between and within countries. The Melanesian countries have a high average elevation above sea level, but also include low-lying islands and atolls such as Bellona raised atoll and Ontong Java in the Solomon Islands, the Lau group in Fiji, and hundreds of small islands and atolls in Papua New Guinea. Apart from Nauru and Niue (countries that are made up of only one island) many PICs are comprised of many inhabited islands, resulting in large Exclusive Economic Zones (EEZs). The distance between islands, both between and within countries, presents a major challenge with respect to trade, transportation, communication, and infrastructure development (Sisifa et al. 2016).

Geopolitically, the Pacific Islands were under the colonial rules of various powers such as Great Britain (administered by New Zealand and Australia), Germany, France and the USA. Currently, PICs are made up of independent states, self-governing states with free association with New Zealand (NZ), unincorporated territories and associated states of the USA and French territories. Sub-regional intergovernmental groups existing in the Pacific include the Melanesian Spearhead Group (MSG) (Fiji, PNG, Solomon Islands, Vanuatu and Kanak and Socialist National Liberation Front of New Caledonia), the Polynesian Leaders Group (PLG) (Samoa, Tonga, Tuvalu, Cook Islands, Niue, American Samoa, Tokelau and French Polynesia) and the Micronesia Presidents' Summit (Palau, Republic of the Marshall Islands (RMI) and Federated States of Micronesia (FSM) (Sisifa et al. 2016). There are also Pacific regional organisations established to provide technical, capacity and operational support to PICs. The key members of the Council of Regional Organizations in the Pacific (CROP) are the Pacific Community (SPC), the Forum Fisheries Agencies (FFA), the South Pacific Regional Environment Program (SPREP), the Pacific Islands Forum Secretariat (PIFS), the South Pacific Tourism Organization (SPTO) and the University of the South Pacific (USP) (DFAT n.d.).

The traditional development partners of the Pacific are Australia, New Zealand, the European Union, the United Kingdom, Japan, the USA and Canada. However,

there is strong emerging support from China, Russia and many other Asian, South American and Middle Eastern countries. New Pacific development alliances are causing discomfort to the traditional partners (Wallis 2017). Non-governmental, development banks and inter-governmental partners are also operating in PICs through several modalities including bilateral (e.g., national governments sectors and budgetary support, non-governmental organizations (NGOs), community-based) and multilateral arrangements, using different implementing agencies (e.g., UN agencies, CROP agencies, and Big Non-Government Organizations (BINGOs). PICs through their own initiative, are also implementing complementary activities to promote sustainable development at the national and sub-national levels (Uitto et al. 2017) and transferring learnings from evaluations to other projects.

Current process of evaluation in the Caribbean and Pacific

The first issue faced by evaluators working in the two regions is the scarcity of M&E systems whether at the project level or nationally. For PICs, lack of impact and outcome evaluation is possibly due to limited demand from governments and high costs involved with access to rural/inter-island places where projects are being implemented (Swan 2016). Effective implementation of sustainable development relies on effective M&E for tracking progress towards desired changes (Uitto 2017). As the PICs strive to achieve sustainable development, a good M&E process is crucial for providing accountability to donors, development partners and communities, and facilitating a learning process so that successes can be replicated, and mistakes can be reduced and avoided. Identifying the reasons why things did or did not work, including reflecting on the specific economic, social and environmental context and to what extent lessons are generalisable, is quite important (Cardno 2017). Tracking progress would not only require M&E systems to be established. It would also require that decision-makers demand and use reports based on systematically collected and analysed quantitative and qualitative data, to revise operations and apply lessons learned to new projects, and not rely solely on anecdotal information.

Like in the PICs, M&E systems in the Caribbean region also tend to be quite weak, and in some cases non-existent, adversely impacting evaluators who work in the region. Funders, borrowers, and evaluators debated this during the 2016 symposium 'Strengthening the Role of Evaluation in the Caribbean Region: Lessons from the Field', which was a collaborative venture with the Office of Independent Evaluation at the Caribbean Development Bank (CDB), The University of the West Indies (UWI), and the Carleton University in Canada. The following issues were

highlighted during the discussions. The general paucity of data in the region is exacerbated by incorrect indicators being used to collect information, non-existent baseline data, insufficient data collection overall, and where data is collected, it is frequently of poor quality, not timely, and inaccessible because of how it is organised and stored. The high public sector turnover rate and a general lack of a public sector evaluative culture (Persaud and Dagher, forthcoming) also inhibit the access to relevant data needed to produce reliable and valid evaluations. Thus, while a data revolution, and by extension M&E, is key to moving towards sustainable development (A World that Counts 2014; United Nations 2016) of the region, high turnover rates caused by fiscal adjustments are negatively impacting efforts in M&E training, and will directly impact the achievement of the SDGs (Persaud, in press).

Evaluators working in the two regions face a second common issue of who leads the evaluation process and the competence of the evaluation team. Evaluation projects in the PICs and the Caribbean have always been driven by donor agencies and mostly performed by evaluators external to the region. The understanding of the required data for M&E has been gleaned through the narrow lens of results frameworks designed for individual projects supported by various funding agencies. Many of these agencies implement M&E training for project personnel, public officials, civil society and grantees engaged in various national and regional projects to demystify what evaluation data should comprise, but the training is often geared towards the implementation and completion of specific projects. Further, language barriers (especially within the Caribbean) have impacted the ability of regional persons to learn from one another and seek and access M&E training opportunities within the region. Another issue is that the training that comes from institutions in the Global North is frequently perceived to be of a higher quality and value than training from local or regional actors. This was the impetus for the formation of the Caribbean Evaluators International (CEI), as a professional organisation of evaluation practitioners of varying competencies who are primarily from the region and work in the region. The initial mission of the CEI focused on building capacity, advocating for evaluation practice, and the recognition and utilisation of regional professionals who have been trained in the field of evaluation beyond what is presented in results frameworks.

A crucial challenge with evaluation in PICs is the limited existing in-country technical capacity to evaluate projects and programs (Uitto et al. 2017). Limited capacity is due to the small population size and "brain drain" (the phenomenon of people with needed skills and qualifications moving overseas in search of jobs with higher salaries). Countries with small numbers of people struggle to collect information and report against all indicators needed by external agencies. The most capable

M&E people have been recruited by regional and international agencies to work on multi-national projects. At the national level, most project coordinators and M&E officers are technical people (e.g. accountants, engineers, medical doctors), sometimes with some level of project management experience – but who lack in-depth knowledge about M&E processes. Development projects tend to have varying indicators, which demand different types of data. Therefore, local project staff is often not clear on what M&E information to collect and how to collect them. There are limited tertiary education courses focused on M&E processes and evaluation methodologies. Most M&E practitioners learn the processes by doing them. There is little existing support in the PICs apart from workshops hosted by regional organizations or project implementing partners to develop evaluation frameworks or train M&E officers. There is no specific organization focusing on building capacity and promoting impact evaluation in the PICs. Capacity issues also apply to donors and implementing agencies. Often donors and implementing agencies employ few technical people to look after many PICs and they are not located in-country but in centralised countries in the region – making travel for field visits, trainings, and technical backstopping very expensive, short and difficult to plan.

The third common issue is the perception that M&E processes are for donor agencies, or to satisfy funding conditions, but are not of use-value to national stakeholders or national development processes. Currently, most M&E that occurs in the Pacific Islands is focused on donor-funded projects or programs, either midway through implementation or upon completion. The work is usually undertaken by external consultants, and is based on the aims, objectives and planned outputs outlined in the original project design documents. Generally, the log frames create a highly structured framework for evaluation, against which successes are measured. These successes focus on outputs, the timeliness and cost-effectiveness of their delivery and is done near the end of the project, rather than focusing on real and lasting outcomes, after the project implementation period.

In the PICs, among stakeholders, there is a passive approach and negative view of evaluation as punishment especially where it is requested when development projects fail or are the subject of criticism. Relationships are crucial to the culture and politics of the Pacific (Denney 2018). People are closely related in the Pacific communities, which tends to make people less critical about development projects, as criticism may be perceived as being ungrateful to the government and donors that funded projects, or disrespectful to the "cousins" who managed the project. The evaluation reports produced are often shared only with project or program partners and are filed away (either in hard copy or electronically).

Access to, and the priority given to stakeholders' views is the fourth common issue in the regions. In the PICs, evaluations have been undertaken through workshops or interviews with "key stakeholders" and project and program "beneficiaries". There are three key weaknesses of this method: i) Pacific communities have strong hierarchical structures, elevating the voices of the elite in workshops. The people who are selected to represent communities are often those with high status in communities such as traditional chiefs, people with religious positions and those working in the government or NGOs; ii) the people selected by project coordinators and managers to be interviewed by evaluators are project beneficiaries, who are therefore more likely to have a positive view of the project. People who were excluded from the benefits of the project, or are critics, are often not included as part of sampling interviews for evaluation; iii) Pacific communities are highly connected, concerned about reputation, proud and optimistic, especially in front of "foreigners" (Uitto et al. 2017). This optimism sometimes creates a false sense of success, as reflected in M&E reports (UNESCO 2017).

In the Caribbean, there isn't a similar hierarchical structure as in the Pacific for stakeholders. The usual procedure is for different categories of stakeholders (primary, secondary, key, external) not to be in the same session for data collection for an evaluation, and persons generally tend to be frank in their critiques of what they perceive to be the foremost issues. The culture of many national and regional institutions is decidedly averse to criticism of any kind, regardless of how constructive these may be. The highlighting of shortcomings is usually met with push back and strong elements of defensiveness among leadership within these agencies, and evaluators are frequently under pressure to modify comments to satisfy commissioning officers, and senior project managers, even where there is an agreement of the rank and file officers with the evaluation findings. The fear is that low scores on aspects of implementation, regardless of the causal factors, may mean that the funding agency will not be willing to provide future funding. The general focus of current donor driven evaluations is to provide information primarily for the funding agency, with limited attention to the crucial aspect of learning and improving within the implementation context, which would in turn be beneficial to implementing agencies and the target beneficiaries. While current Development Assistance Criteria (DAC) guidelines and evaluation guidelines by many donors give lip service to consultation with beneficiaries, the mandate to report back to beneficiaries and local implementing parties is very weak in most Terms of Reference documents, if present at all, and the decision to fully share the M&E reports with these agencies is left to the discretion of the project managers within the funding agencies.

The fifth issue for the regions arises from the conflation of the processes of project evaluation, with those of impact and outcome evaluations. M&E project reports can provide pertinent information about: (1) the short-term progress made that can inform adjustment to processes during the implementation phase, (2) factors which facilitated and hindered implementation; (3) what is already established which can be leveraged to improve implementation; and (4) what may be sustainable without project funds. Impact and outcome evaluations can provide insight about behavioural and systemic changes, but they require a longer time for stakeholder engagement and observation in the implementation context to ascertain the achievement of outcomes. The timelines for most evaluations do not allow for a lengthy engagement by evaluators. There is a heavy focus on the log frame elements and not the theory of change with project evaluation. In the PICs, the need to conduct both M&E and impact and outcome evaluations is crucial, not only to improve processes but more importantly to evaluate the long-term impacts of interventions on behavior, living standards and the resilience of households.

The sixth issue encountered by the regions concerns the dynamics in the relationship with funders, which pit accountability, feasibility and quality against what is practical, relevant and sustainable post-implementation. Many times, the theory of change and its relationship to the logical framework is not documented or clearly articulated to project personnel and relevant stakeholders. The indicators may be poorly defined or inappropriate because the officers in charge of review and quality control do not have the capacity, training, and/or time to make changes to suit the context. Some agencies have a standard set of indicators which countries must agree to measure, which can be problematic for the countries but necessary for the funding agency to facilitate comparisons. External M&E processes sometimes require that evaluators develop a theory of change (TOC) as well as an updated or repurposed log frame with redefined indicators against which to monitor or evaluate progress. The mismatch between high-level, complicated M&E frameworks and country-specific capacity could be easily addressed when donors and development partners consult more closely with in-country partners about what is realistic on the ground before finalizing the process (Dean et al. 2016).

The financing of M&E is a challenge for both Pacific and Caribbean countries. Across both regions, M&E usually takes a back seat to country priorities because the money needed for M&E is viewed as competing for scarce financial resources, which could be better utilised elsewhere.

Transformation priorities for development

The climate change vulnerability of the two regions exacerbates the systemic weaknesses in both regions, to access and sustainably use resources for food security, housing, healthcare, livelihoods and maintain security systems. The Caribbean continues to face the degradation of its ecosystems and environment, the plundering of its natural resources due to inefficient consumption and production patterns, and the increasing vulnerability of its population due to the global challenge of climate change (ECLAC 2018). The hurricane season of 2017 brought devastation to Dominica, the British and US Virgin Islands, Puerto Rico, Sint Maarten, the Bahamas, Cuba, the Dominican Republic, Haiti, Anguilla, Martinique and Guadeloupe. While the regional response mechanism was activated by the Caribbean Disaster Emergency Management Agency, and aid agencies from the US, UK and countries external to the region, the disasters highlighted the vulnerability of the region and the magnitude of the financial resources needed for recovery.

The Pacific region is faced with adverse impacts of human induced climate change. The sea level is rising, causing forced migration, coastal erosion, saltwater inundation and intrusion affecting coastal communities' water and food security (Iese et al. 2018). The PICs have very high exposure to extreme events influenced by climate change (e.g., tropical cyclones, floods, storm surges, landslides, droughts) and non-climatic extreme events (e.g., volcanoes, earthquakes, tsunamis) (Fepuleai et al. 2017). Furthermore, the high incidence of non-communicable diseases (NCDs) and the fast rate of urbanisation are immediate threats to Pacific communities and environmental sustainability (Iese et al. 2018).

Pacific island leaders have repeatedly emphasised the urgency of addressing climate change in regional fora, through the implementation of various global and regional agreements and frameworks. These include the UN Paris Agreement, the Boe Declaration and the regional Framework for Resilient Development in the Pacific (FRDP). Furthermore, Pacific Island leaders have committed to realising the 2030 Sustainable Development Agenda and the SIDS Accelerated Modalities of Action (S.A.M.O.A) Pathway, which relates to the SDGs. Other priorities such as the long-term sustainability of fisheries resources, the fight against non-communicable diseases, and the promotion of culture are crucial for building a strong Pacific region and addressing the challenges outlined above (PIFS 2018).

The Thomas and Benjamin study (2018) also brought to the forefront the issue that climate change may cause forced migration with increased loss and damage

to coastal areas, thus the adaptation strategies in the two regions must include planned relocation and migration to avoid citizens from becoming climate refugees. The study further noted that the issue of migration was not addressed in the climate change and disaster risk plans, strategies of policies coming from the two regions. While donor agencies tend to ask for Impact evaluations, rigorous mixed methods approaches can assist with the revision and development of suitable multi-faceted plans, strategies and policies to facilitate migration as an adaptation strategy.

The international community has been challenging SIDS to take ownership and lead their own sustainable development (A World That Counts, 2014). The many challenges involved with financing sustainable development in the Caribbean and Pacific Islands regions are rather complicated since the regions face many vulnerabilities. Compounding this issue is access to concessional international financing which presents an obstacle for many countries because of high public debt and large fiscal deficits, which reduce their fiscal latitude to obtain the financing needed to achieve their SDGs. Specifically, in 2014, the debt-to-GDP ratio of many Caribbean countries exceeded the internationally accepted rate (Hurley 2015). Given these realities, the two regions will need to strategize on how to get dedicated resources to assist with mobilizing progress towards achieving their SDGs.

Transformation and the use of evaluation

The authors recognise that the envisioned transformation of the Caribbean and the Pacific regions, through the implementation of the vision of the UN 2030 Agenda to meet the targets aligned with the SDGs requires a major shift in thinking about national development, the use of resources and the use and value of evaluation. While the team can identify the issues that external agencies must consider when working in these regions, the team also recognises that there must be thought leadership and action leadership within their own regions. The team makes the following suggestions for necessary actions to be taken considering the actors internal and external to their regions.

Evaluation culture The establishment of an evaluation culture is essential to achieving transformation through the SDG goals. Data must be valued, with data collection, analysis, reporting and learning becoming routine operations for governments and not just the work of one person or unit. There must be coordination among line Ministries to avoid data duplication and redundancy and facilitate the improvement of data quality. The opportunities to learn must include

the analysis of unintended consequences of project implementation, reviewing negative feedback without defensiveness, and acknowledging that adjustments to procedures will improve progress.

Leadership Buy-in at the leadership level of government is needed to support and authorise the development and establishment of policies, protocols, procedures and the standardization of data collection tools that will facilitate comparability of data when necessary and the efficient use of an M&E system. For transformation at the national level, public sector organizational policies must mandate M&E as a requirement and commit to M&E through continuous training of personnel. The strength of the commitment can be demonstrated by each country's government assigning a budget for M&E in the annual government estimates of revenue and expenditure, and leading M&E initiatives to encourage public sector employees to buy-in to the process.

Role of the evaluator Transformation will also require substantial work to demystify evaluation and the role of the evaluator so that an evaluative culture can truly develop. This is needed to overcome the current fear of evaluation, which is exhibited by anxiety behaviours such as stakeholders' unavailability for meetings, data and documents being promised but never shared, and only the top official in a team speaking in meetings. Evaluators should guide stakeholders away from fear and blame, so that learning and improvement become strategic actions. Helping stakeholders to see the big picture of how each successful program can contribute to the government's overall strategic plan can also help to put stakeholders at ease (Persaud, in press).

Data collection For evaluation to help with transformation, national and regional agencies must be established to collect SDG implementation data. Many countries do not have written plans which outline a systematic strategy to sensitise relevant stakeholders about the SDG goals and the achievement of targets. National reports should be informed by procedures which document the in-country work of donor agencies, NGOs, Ministries, or any organisation which is undertaking projects focused on the SDGs.

Evaluator competencies The improvement of evaluation utility and use will also be contingent on building trust and buy-in to evaluations. In this respect, evaluators should demonstrate sensitivity to various stakeholders' concerns. There is a strong need to build the capacity of local evaluators who understand the languages, culture and relationships within countries in the Pacific and Caribbean regions. For evaluation to have a transformative impact, professionalisation of evaluation must

be considered – the knowledge of results frameworks will not be enough. The study of different evaluation approaches (e.g. utilization-focused, development, developmental, empowerment) is necessary given the context and purpose of the evaluation not the purpose of the project, especially as sustainability initiatives must be the responsibility of governments, not funding agencies. Thus, there is a role for partnerships within and between regional and national tertiary institutions and governments, leveraging the relationships not only to conduct training for public sector personnel but also for research and knowledge dissemination. The inclusion of trained national and regional M&E practitioners in the design phase will enhance the transformation process as they share their insight of the implementation context, in the crafting of theories of changes (TOCs), and choices of indicators and outcomes.

Reports A repository of Evaluation Reports commissioned by international development partners should be readily available in an accessible online location for the availability of all interested parties. The reports should be searchable by keywords, themes, geographic location, topics, timelines, and issues addressed.

Sharing lessons learned While monitoring and especially evaluation reporting frameworks frequently require a lesson learned section, there is no overt intention to share these lessons with anyone other than the immediate project personnel. Many times, the wealth of knowledge that evaluators gather from seeing multiple projects across funding agencies with the same recurring issues, has not been shared and confidentiality clauses limit what can be shared. Transformation must involve opportunities for practicing evaluators to share their experiences by writing articles, participating in conferences, and encouraging such discussions in their professional organisations.

Conclusion

The small island developing states in the Caribbean and the Pacific have a history of collaborative efforts to address economic development. The vulnerability of the two regions has increased in the wake of climate change patterns which have illuminated development gaps in the areas of food security, housing, health, education and citizen security for their populations. The targets of the sustainable development goals indicate pathways which can be followed to transform the regions. However, the two regions are still grappling to prioritise the use of investment resources for their national development and still comply with the UN 2030 Agenda.

The authors discussed how evaluations could be leveraged to help regional countries make progress towards development outcomes while addressing the SDG goals. However, they were also able to identify the issues which diminish the utility and quality of evaluation work. They identified the weak M&E systems, leadership of evaluations, the perceived low technical capacity of regional personnel to conduct evaluations, access to different types of stakeholders, project versus impact and outcome evaluation, and the financing of M&E activities as the major factors that negatively impact the value of evaluations conducted in their regions. The building of an evaluation culture, garnering political will at the highest levels of government, recognising the critical role of evaluators, identifying desired competencies for evaluators and sharing the lessons that would facilitate transformation of the regions, are the solutions they proffer at this time.

The Caribbean and the Pacific regions remain heavily dependent on donor funding and investment to realise their national development goals. Thus, evaluators external to these regions will continue to be engaged to do the required M&E, so efforts to organize and undertake indigenous evaluation to open up perspectives that would be closed to donor related evaluation practices have to be embraced. The transformative value of evaluation lies in the benefits that countries in the Caribbean and the Pacific derive from evaluation findings, that would guide their development planning to access and use resources to reduce the inequities between themselves and the more developed countries in areas such as climate change, education and health, but that shift in thinking of regional leaders must come first.

References

CARICOM, and SPREP. 2004. *Adapting to a changing climate in the Caribbean and South Pacific regions. Guide to the integration of climate change adaptation into the environmental impact assessment (EIA) process*. [Georgetown, Guyana:] Caribbean Community (CARICOM) and South Pacific Regional Environment Programme (SPREP). https://tinyurl.com/y45fsf9p.

Cardno. 2017. Monitoring, evaluation and learning framework. *Pacific women shaping Pacific development.* Cardno Emerging Markets. https://tinyurl.com/yxkzngny.

Dean, Annika, Donna Green, and Patrick Nunn. 2016. Too much sail for a small craft? Donor requirements, scale and capacity discourses in Kiribati. In: *Island geographies, essays and conversations*, (ed.) E. Stratford. New York and London: Routledge.

Denney, Lisa. 2018. *Walking the adaptive talk*. https://tinyurl.com/yxkn23sp.

Dados, Nour, and Raewyn Connell. 2012. The Global South. *Contexts* 11 (1): 12-13. https://journals.sagepub.com/doi/pdf/10.1177/1536504212436479

ECLAC Economic Commission for Latin America and the Caribbean. 2018. *Access to information, participation and justice in environmental matters in Latin America and the Caribbean. Towards achievement of the 2030 Agenda for Sustainable Development* (LC/TS.2017/83). [Santiago: ECLAC]. https://tinyurl.com/y5jhsl92.

Fepuleai, Aleni, Eberhard Weber, Karoly Németh, et al. 2017. Eruption styles of Samoan volcanoes represented in tattooing, language and cultural activities of the indigenous people. *Geoheritage* 9 (3): 395-411. https://doi.org/10.1007/s12371-016-0204-1

Galloway McLean, Kirsty, and Serena Heckle (eds). 2017. *Towards Climate Change Resilience Minimising Loss & Damage*. Paris, and Apia, Samoa: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000260140.

Hurley, Gail. 2015. *Financing for development and Small Island Developing States: A snapshot and ways forward*. New York, UNDP and UN OHRLLS Discussion Paper. https//tinyurl.com/y5au8fsl.

Iese, Viliamu, Elisabeth Holland, Morgan Wairiua et al. 2018. Facing food security risks. The rise and rise of the sweet potato in the Pacific Islands. *Global Food Security* (18): 48–56. https://doi.org/10.1016/j.gfs.2018.07.004.

Persaud, Nadini, and Ruby Dagher. 2019. Evaluation in the English-speaking Commonwealth Caribbean: lessons from the field. *American Journal of Evaluation*. https://doi.org/10.1177/1098214019866260.

PIFS Pacific Islands Forum Secretariat. 2018. *First Quadrennial Pacific Sustainable Development Report*. Prepared by the Pacific Islands Forum countries with the support of the CROP and UN agencies. Suva, Fiji, Council of Regional Organizations in the Pacific, UN Agencies in the Pacific.

Persaud, Nadini. *Strengthening programme evaluation culture in the English-speaking Commonwealth Caribbean. A guide for evaluation practitioners and public sector decision-makers*. Kingston, Jamaica: Arawak Publications (in press).

Robinson, M. 2014. The debt experience of SIDS in the pacific. In: *Debt and development in small island developing states*, (eds) Damien King and David F. Tennant, 29-70. New York: Palgrave.

Sisifa, Aleki, Mary Taylor, Andrew McGregor et al. 2016. Pacific communities, agriculture and climate change. In: *Vulnerability of Pacific Island agriculture and forestry to climate change*, (eds) Mary Taylor, Andrew McGregor and Brian Dawson, 5-45. Noumea, New Caledonia: Pacific Community (SPC). https://tinyurl.com/yyaue9s9.

Thomas, Adelle, and Lisa Benjamin. 2018. Policies and mechanisms to address climate-induced migration and displacement in Pacific and Caribbean small island developing states. *International Journal of Climate Change Strategies and Management* (1): 86-104. https://www.emeraldinsight.com/doi/pdfplus/10.1108/IJCCSM-03-2017-0055.

Uitto, Juha I., Jeremy Kohlitz, and David Todd. 2017. Evaluating sustainable development in small island developing states. Lessons from the Pacific and the Caribbean. In: *Evaluation for Agenda 2030. Providing evidence on progress and sustainability*, (eds) Rob D. van den Berg, Indran Naidoo, and Susan D. Tamondong, 119-133. Exeter, UK: IDEAS. https://ideas-global.org/evaluation-for-agenda-2030/.

United Nations DESA/Population Division. 2018. *World urbanization prospects 2018* [dataset]. https://esa.un.org/unpd/wup/download

A world that counts. Mobilising the data revolution for sustainable development. 2014. New York: United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG). https://tinyurl.com/lx8mj7d.

Value-based evaluations for transformative change

Inga-Lill Aronsson and Hur Hassnain

ABSTRACT. This essay argues for "value-based evaluation" for transformative change. It emphasizes the transformative power of inherent, heritage-driven values in fragile and conflict-affected contexts, in which they are often targeted and destroyed, and points out the lack of reference to them in the SDG 16 and the Agenda 2030 as a whole. The authors indicate that in conflict-sensitive evaluations (heritage-driven) values can play a significant role as connectors, in building connections with stakeholders. The authors present suggestions from their experience working in some of the worst conflict-affected countries in the world, and at the same time acknowledge that it is a highly delicate task for an external evaluator to bring out local, institutional and national values even when they can be linked to sustainable change. They recognize the double-edged nature of heritage in conflict sensitive environments.

Introduction

This essay is about how to use value-based evaluation for transformative change especially in countries torn by humanmade and natural disasters. A value-based evaluation is an approach based on primarily, but not exclusively, the values of evaluand and the subject of evaluation. Value-based evaluation is more than participation and cultural sensitivity. It is an understanding of deeply rooted values inside the community or an institution that the evaluator needs to articulate in a participatory way and find methods for articulation and theories and methods that support this articulation. The value-based evaluation aims to look deeper into the socio-cultural and economic-political constructs relating to the local and institutional values that the evaluand articulates especially in a conflict-stricken environment to suggest a transformative change through conflict transformation.

In fragile, and conflict-affected settings, the value-based evaluation looks specifically into what individuals and formal as well as non-formal institutions value, that later build into dividers or connectors of or in conflict, looking specifically into some of the underlying issues and related concepts such as "power" and "subjectivity". These values are deeply connected to the community or institutional heritage that also defines ethical standards for the whole evaluation cycle, starting from its design, implementation, communication and uptake. Identifying such values could also help an evaluation to develop and/or refine its instruments/tools, such as using the power of music, dance, stories and art that can also be target in times of conflict. The world has also seen people sacrificing everything for their heritage and values, for example, Curator Khaled al-Asaad died to protect the history he loved. When ISIS fighters took control of the historic site of Palmyra, Syria, in 2015 and demanded him to tell them the location of valuable items, he refused, and was murdered. An evaluator in a country torn apart by conflict, such as Syria, has a greater responsibility to identify and include diverse, multiple and under-represented perspectives to help achieve sustainable peace. Understanding heritage driven values helps an evaluator understanding different perspectives and the dividers and connectors of conflict and violence.

The case example taken in this essay focuses on heritage driven values in Sierra Leone; a state in non-linear transformation, which has been through episodes of unprecedented, but also varied levels of violence, civil unrest, natural disasters and Ebola in just a few decades after its birth in 1961 as an independent state. Evidence highlights that the Sierra Leonean society continues to struggle to tackle its poverty levels, weak GDP growth and Human Development Indicators due to its rapidly shifting context. How can one evaluate in such a fluid and unpredictable

environment? What can be the underlying values associated with social change in such a society, and what are the key challenges to look at while designing, conducting or communicating an evaluation if the context is so unpredictable, conflict-affected and complex?

This essay argues on its most general level that if heritage is a contributory factor in causes of conflict, it might also be used for uniting people in a national narrative. An evaluator must put heritage driven values in the center of the whole evaluation cycle so it can identify and work on and through the connectors and separators of drivers of conflict and violence.

A theoretical blind spot in global thinking

Heritage is always ambiguous and symbolic, tangible and intangible. For some people, some events and sites represent greatness, for others pain and suffering. Therefore, heritage in conflict is always contested and exposed to targeted destruction. Monuments are being torn down, relocated to a monument graveyard, or simply destroyed in order to get rid of a past that is no longer valued. For example, the Buddha statues in Bamiyan in Afghanistan were destroyed by the Taliban in 2001. The Taliban even announced the destruction beforehand, to gain maximum media attention, which it also got, with a global condemnation as a result.

Heritage can be destroyed and then rebuilt to "heal" a society after conflict. For example, the Mostar Bridge in Bosnia-Herzegovina was targeted and destroyed in 1993 during the Balkan wars. The bridge was an Ottoman architectural masterpiece, built in the 16th century. UNESCO decided to rebuild the bridge as authentic as possible, using the original stones, the original architectural plans and as far as possible the original construction techniques. Authenticity counted here. It was re-opened in 2004 and declared a UNESCO world heritage site in 2005. The cost was more than 15 million US dollars. The bridge was rebuilt to reconcile the divided city. This kind of reasoning's underlying logic is that it would trigger an almost mystical transformation process between the built environment and the people (Aronsson 2013). However, critical voices were heard in the local society; the money could have also been spent on schools and on infrastructure between urban and rural areas to improve services and trade.

To return to the Buddha statues, before the destruction, they were a global tourist attraction. After the destruction, the tourists came to visit the empty hole in the mountain. Lately the destroyed statues have been "reconstructed" with 3D light

projection techniques for tourists. Hence, at first, the authenticity of the statues does not seem to matter here for the tourists, and it could be argued that they only come to experience the destruction of the site, which is known as atrocity tourism. This might be so, but it is a simplification of the complex relationship between people and the material heritage. Authenticity matters, but it consists of several layers and relations. In this case, the gone material statues and the present site. Peoples´ gaze changes and they imagine the past site, the destruction and the future. It makes people think. The merging of temporal and spatial orders is a vital capacity of heritage. For Taliban the statues were idols, (which goes against their religious values) but they worked hard for 25 days to blow them up. When the world objected (because of its unique value for humankind), they reacted with blankness and said that they were only destroying a heap of stones.[1]

There are abundant similar cases throughout the world to discuss, similar at least on a superficial level. Nevertheless, tangible and intangible heritage is resilient and transformative at the same time. How can an evaluator, who is often an outsider, understand the underlying values associated with the heritage and build on those to identify and report on what actually matters to the people, what has actually changed over a period of time, because of what, for whom and how? Is it even desirable to get so close to another culture? Does it belong to the work description of an evaluator? Those questions are formidable challenges for evaluators and the entire field of evaluation especially in the context of transformation in fragile societies. The underlying assumptions are that it is possible to distinguish between outside and inside knowledge, and to objectify events and people's life-worlds and livelihood and correlate them to a comprehensive evaluation scheme with predefined targets and indicators. Anthropologists have been struggling for almost a hundred years with who-is-who, and what-is-what in ethnographic accounts: What kind of knowledge is produced? Who is represented, and re-represented in the ethnographies? Emic (inside) and etic (outside) knowledge, mixed teams, subcultures/subgroups, asymmetrical power relationships are just the beginning. On top of this is the time dimension with evaluators working under harsh deadlines. On the other hand, what are the options? This is where we are today, and, in this essay, we introduce heritage driven values in evaluation as one possible path to explore, which makes it even more complex.

In order to evaluate interventions in a complex, fluid and volatile context, the first

1 See AFP. 26 February 2001. Afghan leaders order destruction of ancient statues. http://www.rawa.org/statues.htm.

task of evaluators is to have a comprehensive and holistic understanding of the context especially its underlying heritage-driven values and their interplay with the drivers of violence. "In Southern Sudan for example, it was found that the support provided by multiple donors in 2005-2010 was often mistargeted, because donors did not fully take into account key drivers of violence, there was an overemphasis on basic services and a relative neglect of security, policing, and the rule of law, which were found to be essential in the process of state formation for the future South Sudan and therefore, critical to preventing future conflict" (Bennet et al. 2010).

This essay will look closer at Sierra Leone from the above-presented lens. What are the values Sierra Leone holds as represented in the heritage sector and how can an evaluator tap on those values to evaluate a programme and come up with applicable recommendations that suggest a sustainable transformative change?

The United Nations Sustainable Development Goal (SDG) 16 states, "Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels".

SDG 16 has 12 targets with several indicators each, the highest number of indicators amongst all the SDGs. Notwithstanding its comprehensiveness, the only target that might somehow refer to heritage is target 16.A, which highlights the strengthening of relevant national institutions, to prevent violence and combat terrorism and crime with its associated indicator of human rights.[2]

The lack of explicit consideration of heritage and memories as building blocks in the transformation of fragile societies is a weakness of SDG 16. It is a theoretical blind spot in the agenda 2030, which might explain why we hardly see the ability to report outcomes so far. We suggest that the powerful discourse of SDG in its present form partly determines and informs research and evaluation projects, which leads to the risk that the mentioned building blocks of heritage and memories are not sufficiently considered in a transformative evaluation context.

2 See https://sustainabledevelopment.un.org/sdg16

Sierra Leone: a country in the Global South

Conceptually, the Global South is associated with previous concepts such as "Third World", and "developing world". Furthermore, The Global South "refers to countries concentrated around the equator and in the southern hemisphere with relatively low levels of both geopolitical power and development..." (Leibfried et al. 2015, 2).

The prevalent view of Sierra Leone is that it is a country in the Global South with a colonial history, internal conflictual past with an aggregation of local wars, ethnic identity crisis and instability as a national state. According to the Organisation for Economic Cooperation and Development (OECD), Sierra Leone stands at number 33 out of 58 states in the Fragile State Index 2018 (OECD 2018).

The slave trade started in the 15th century. On March 1807, the Slave Trade Act, officially an Act for the abolition of the Slave Trade, entered the books of the United Kingdom (UK) prohibiting the slave trade in the British Empire. Nevertheless, trafficking continued until several decades later in the world, until the markets for slaves in the Americans and the Caribbean disappeared.

The capital city of Sierra Leone, Freetown, was founded in 1787 by the Sierra Leone Company (SLC). Around the same time, British Naval Lieutenant John Clarkson brought back freed American slaves from Nova Scotia, Canada to Freetown (Banton 1969). Hence, former slaves were shipped to Sierra Leone after generations as slaves, a country they had no knowledge of and belonging to. From 1807, when the UK abolished slave trade, slaves freed by the UK from slave vessels of other countries were also disembarked in Freetown. This provided part of the new population of Sierra Leone.

Sierra Leone became independent from Great Britain 1961. Four years later in 1965, the first and so far the only inventory of cultural heritage sites was implemented (Basu 2014, 236). Most of the inventory took place during the colonial rule.

The eleven-year civil war in Sierra Leone lasted between 1991 to 2002 and resulted in the death of over 50,000 people. The main actors of the civil war were the Revolutionary United Front (RUF) against the Sierra Leone Army (SLA), but many more militant groups were involved fighting each other, with diverse and loose loyalties. The war also involved conflict diamonds and created a "path dependence" that continues to hold back the development in Sierra Leone till date.

Not long after years of brutal violence, Sierra Leone fell into the hands of the Ebola virus epidemic in 2014 with severe consequences for the entire society. Thousands of people died, and this had a major impact throughout society and the economy. Since then, the country has fallen more deeply into aid-dependency, institutionalized corruption, and much more (UNDP 2016; Transparency International 2019).

Needless to say, Sierra Leone has been ill-fated since it came to birth. Have we experienced any examples of societies bouncing back from all these atrocities? How have these atrocities affected Sierra Leoneans´ mental, emotional, social and physical resilience? Have societies become more absorptive, adaptive or transformative in nature due to experiencing these shifting environments? If yes, how can an evaluator tap on such questions in a context like Sierra Leone? Has heritage anything to do with it?

On the one hand, according to Basu (2014), the country has to "confront" its past to gain an inclusive national narrative, which would unite the people and support a collective identity. The neglect of the past only obstructs the transformational healing processes (ibid 2014, 233-234). On the other hand, maybe the country bounces back, because it has not yet confronted its past and opened its Pandora box of the past. In the research on reconciliation, heritage and resilience this is still an open question. The choice not to deal openly with a difficult and dark past through mechanisms such as truth commissions, the digging up of mass graves and peace projects is rather common for post-conflict states. For example, in Spain, Franco followed his victory in the civil war in the 1930s with oppression. After his death in 1975, the identification of mass graves became a folk movement; people demanded justice and legal consequences for the perpetrators. This is one argument for taking heritage and memory seriously in evaluations, but we still do not have sufficient empirical research data to claim that it is better to confront the past, than to silence it.

Furthermore, it could be argued, that to truly transform into a sustainable (in all its dimensions) state Sierra Leone needs to take into consideration its heritage in all its diversity, tangible and intangible, dark and difficult, but also rich and beautiful. But then it depends on what is meant by a real transformation to reach sustainability. It is probably not a transformation that is based on a "Disney" like reconciliation with a violent past.

Heritage in Sierra Leone

The independence struggles were not followed by "a significant cultural nationalist movement and thus the museum and other cultural institutions lacked an important raison d'être. It could be argued that, as yet, they still have not found one" (Basu 2014, 234-235). If such a national movement would have supported reconciliation is an open question in this diverse society. The relations between heritage and ethnic identity is mostly taken for granted in the heritage literature, albeit without a sufficient problematization of either concepts or presentation of ethnographic material that would empirically show how this relationship is articulated on the ground.

Why a significant national movement did not take place, might partly be explained by the country's colonial past and the way in which the indigenous heritage was collected and displayed in museums all over the world, if we accept the premises that heritage is a kind of binding glue and that material objects represent groups of people. Theoretically, when a material object (e.g. a ritual object loaded with symbolic significance) is selected for a museum collection, the object is transformed in its journey through several cultural domains before it reaches the museum (Appadurai 1986). Basically, it is removed from its socio-cultural context in the real world (cleaned, investigated, numbered, classified, related to other objects in the same categories, displayed) into the museum context. The objects' primarily purpose becomes to generate knowledge; it is transferred and arranged into a predefined classificatory order of the museum. The object displayed in a glass showcase loses its original purpose, and in the visitor's gaze it is transformed into something of the past together with the people who owned and used it. Hence, the colonial powers undermined on a very deep level, the value-based heritage, and made a comprehensive national narrative more difficult to articulate by collecting every day and ritual objects from groups of people and turned them into museum objects.

The colonial powers collected the ritual and religious objects with the aim of destroying, or transforming, the local culture and its symbolic system in order to convert people to Christianity (see Okeke et al. (2017) for an elaborate discussion). The objectification of people's life worlds always undermines the socio-cultural structures of the exposed society. This goes for past practices, as well as for large infrastructure development projects of today (Aronsson 2002). This does not mean that a local heritage in Sierra Leone is not present, on the contrary; there is a rich and varied heritage throughout the country that is proudly maintained and practiced, but it is probably a challenge to create a meaningful and articulated

narrative out of these fragments. The society is fragmented in its very core.

Nevertheless, transformations take place. For example, the Sierra Leone heritage site presents the heritage sector in the country and its global connections through on-line access to collections in: Sierra Leone National Museum (1618 objects); British Museum (882); Cootje Van Oven Collection (236); Brighton Museum and Art Gallery (254); Glasgow Museums (309); World Museum Liverpool Sierra Leone Collection (182) and British Library (110)[3]. This could be thought of as a minor accomplishment in a country with immense problems, but maybe the old anthropological saying still holds: "the action is in the homestead", where also the transformation slowly takes place.

The webpage also presents the 16 tangible heritage sites (out of originally 18, but two are destroyed) that to date are protected national monuments under Sierra Leone's 1946 Monuments and Relics Ordinance. Out of these 18, only one is not associated with the colonial or slave trade history (Basu 2014, 236).[4]

The last proclamation was made in 1965; hence, the majority was declared a heritage site during the colonial period. It is highlighted that there are many more historical and archaeological sites in the country and that each village has its own tangible and intangible heritage. Furthermore, there is a video gallery with 47 videos presenting each a heritage theme/practice such as the world-famous carved female figures, or the national history of Bai Bureh (a freedom fighter).[5] These videos introduce both tangible and intangible heritage. Here we also discern a transformation in the small format that might have great transformational power in the long run. The heritage sector is funded by Arts & Humanities Research Council, Beyond Text, UCL, University of Sussex, School of Oriental and African Studies - University of London.

---

3 See http://www.sierraleoneheritage.org/sites

4 The heritage sites are: Bunce Island; Heddle's Farm; De Ruyter Stone; Bastions of Fort Thornton; Earthworks and Live Stockade at Masakpaidur; Gateway to the King's Yard; Ruins of the John Newton's House and Slave Barracoons, Plantain Island; Cleveland Tombstone; Old Wharf Steps and Guard House; Old City Boundary Guns; Old Fourah Bay College Building; St Johns Maroon Church; St Charles Church and King's Yard Wall; Firing Point and Guns near Old Wharf, Dublin; Banana Island; Martello Tower; Grave of Captain Lendy, Waiima. http://ww.sierraleoneheritage.org.

5 The videos are: Arong Athoma, Bai Bureh, Balangi, Basket Making, Bee Keeping, Bondo Society, Boui, Carved Female Figures, Chiefly Regalia, Cola Gara, Falui, Fana, Fula Music, Goboi, Gongli, Gumbe, Hunting Society, Jobai, Jobuli, Jollay, Kabemba Mask, Kelene, Kohwaso, Kondi, Kongoli Mask, Mammy Yolo Mask, Matorwa, Mbambue, Mortai, Nafali, Odday Society, Odelay, Society Music, Ojeh, Palm Wine, Pot Making, Pottery, Faffia, Shegbureh, Sowei, Sowei Mask, Stone Sculptures, Stringed Instrument, Tagajesea, Warri, Weaving. http://ww.sierraleoneheritage.org.

Since the ending of the Sierra Leone Civil War in 2002, the international development organizations have turned their efforts to support institutional capacity building of peace and reconciliation. Heritage is one of the sectors that has raised interest, but often as a tourist attraction development industry, which is the case in Sierra Leone. Heritage is seen as a commodity with economic potential. Especially, the slave trade sites and the slave-return tourist industry looking for their roots. This is part of a global trend called "atrocity tourism" or "dark tourism" as mentioned above. There is something cynical here that the destruction of culture during colonial times is used to rebuild (reconcile?) the society of today. Is this exploitation or just a smart way to use the past that anyhow cannot be changed?

What we have here is thus a heritage sector that partly is a reminder of the colonial time, but also a sector in transformation with a widening awareness of the value of incorporating the local indigenous heritage into a national narrative. It is clear that the institutionalized heritage sector in Sierra Leone is transforming in collaboration with the Global North and whose collections (Global South and Global North) are mutually contributing to our knowledge about objects, people and civilizations, made possible with digital on-line search functions in digitized collections wherever they are located physically. Hence, transformation is visible and evaluable.

However, the case of Sierra Leone also shows that its national narrative so far seems to have placed the roots of the conflict outside of Sierra Leone. This is backed up by The Truth Commission Report (2004), which has created a narrative close to "reinventing the past through truth telling". It is an "Anachronistic rewriting of the Sierra Leone's history" that vaguely emphasizes that before colonialism, people lived in "harmonious coexistence" (Basu 2014, 237-239). However, this is not correct. Evidence is available at concrete sites, and is visible and tangible, and stories of past evil deeds are told, danced and sung. Hence, when memories and narratives of heritage are collected from communities around the country, the picture of Sierra Leone emerges as: "Contrary to the narrative of the conflict as an incomprehensible disruption in a long history of peaceful coexistence" we are confronted with "a palimpsest of violence" (Basu 2014, 241).

However, the Truth Commission of Sierra Leone is not alone in its efforts of creating a biased narrative of a dark history. Truth commission reports in general "play a vital role in fixing memory and institutionalizing a view of the past conflict" (Wilson 2001). There is questionable moral discourse of these reports that we find troublesome when considering a value-based evaluation. These kinds of powerful documents and reports often govern evaluation, which means that evaluation can easily reach misleading conclusions and in the long run give faulty policy recommendations.

As already mentioned, we know that a general problem for countries with a dark heritage is to decide when the time is right to bring up the dark past. Is it better to wait until the country has "calmed down", or is it instead better, or maybe even inevitable for progress to confront the past immediately? Here evaluation can contribute to the discussion on transformation with a comparative methodological frame without being wrapped up in high level conceptual reasoning.

Discussion

Social transformation is defined "as the process by which society, organizations, and individual change happens, such as changes in behaviours or cultural norms and perceptions as a direct or indirect result of community action" (Rodriguez-García 2015, 148).

Furthermore, Rodriguez-García, referring to Weiss (1995) suggests that the theory of "change describes the set of assumptions that explain both the small steps that lead to long-term goals and the connections between programme activities and outcomes that occur at each step of the way" (ibid 2015, 147). Consequently, complex programmes are difficult to evaluate. This is often because the assumptions that guide the evaluation are poorly articulated. In cases of heritage driven values, a big question for theory-based evaluation is "whose assumptions?" and "who is assuming?" Theories of Change in such cases may need to be (re) developed at the evaluation stage together with the stakeholders, if needed. This will ensure that values are central to the evaluation design, implementation and reporting. This is particularly important in a conflict struck context, as people have a certain level of belongingness to places, events, rituals, languages and arts and culture.

The involvement of the stakeholders is not free of problems. They can be biased, have their own agendas and moral compasses. They always "read" the evaluator, as much as the evaluator "reads" them. It can be compared with a ritual dance, when both parties have learned how to dance. Evaluation is performance as well as performative.

Value based evaluation is more than participation and cultural sensitivity. It is an understanding of deeply rooted values inside the community that the evaluator together with the people needs to articulate and find methods for articulation and theories and methods that support this articulation. If it is more than local informed participation used in evaluation, so what is it?

It may be the process of an evaluator facilitating people to articulate their being-in-the-world. This facilitating role is tricky but necessary, because people in general

do not express values in words; they rather live those values expressed in everyday practices.

In this process, the evaluator will be queried, as well as the stakeholders. It is naive to assume that people will answer the evaluator´s questions fully. There will always be a front stage and backstage. It is human communication, beyond participation. However, as already asked, how close should an evaluator really come to the people? With what right do we penetrate deep into people's everyday life? Are we not dangerously close to exploitation? Or, is an inter-cultural, open and respectful dialogue possible?

A strong contextual analysis informed by the heritage driven values informs the evaluation's understanding of the sources of tension and cohesion in the communities, the key drivers of conflict, triggers of violence and more importantly the likely future conflict scenarios. A more detailed analysis of what communities value and regard help to understand the background and history of the conflict; identify all relevant groups involved and their perspectives, and the causes of conflict.

Finally, heritage is nationalistic, but in a global world, it is also (or maybe even more) global, as formulated by Cuno in his book *Who owns antiquity?* (2008). In a fragile state such as Sierra Leone it is difficult to say what the implications are for this trend of regarding heritage as both belonging to the country and the world. It depends on where the heritage sector is at the moment in the country's overall transformation. However, evaluation probably has to adapt to this shift and pay attention to international partnerships and reciprocal collaborations.

Social and cultural norms, heritage and transformation are deeply connected. They are intertwined with each other. To bring it even further, norms are culture and heritage (tangible and intangible) and there are no transformative changes that can be induced or evaluated without considering them in evaluation design, implementation, communication and uptake. Norms can be facilitators and/or inhibitors of transformation at the same time. This omnipresence makes it so difficult to operationalize, in order to bridge the gap between value-based heritage evaluation and reality. It becomes even more obscure if one considers the abundant post-modernist literature on heritage that claims that there is "no such thing as heritage" (Smith 2006). Of course, that is nonsense, and we just have to go back to the case of the Buddha statues to see this.

This essay will now elaborate on some suggestions of how a value-based heritage evaluation can be modelled.

- A value-based evaluation must account for a heterogeneity of factors.
- A value-based evaluation in a fragile context must put vulnerability and respect at the centre. However, both "vulnerability" and "respect" can be double-edged. Who is to decide when values clash about what is a good transformation? In other words, the evaluation must have strong ethical foundations.
- The methodologies used must be progressive, culturally sensitive and adopted to the fragile society with its violent past, including contested sites and narratives regardless if it is from colonial times or recent. The narratives of the past play a crucial role in how the society sees its transformation.
- Songs, poetry, handicraft, art, narrative and dance can be studied in order to reach an understanding of its content and role in society. They will not be "assessed" per se, but will be used to understand and come close to the cultural heritage of the country in its richness. There can be ways to articulate these in evaluation tools and adapting/utilizing these to develop participatory evaluation tools and approaches.
- The *institutional* heritage and its associated heritage institutions can be assessed by paying attention to the number and types of institutions (museums, archives and libraries), distribution, accessibility, performance, funding, staff's level of education, collection management, exhibitions and pedagogy.
- The *institutional* heritage sector's international collaborations and exchanges can be assessed by paying attention to type, frequency and funding and its associated power relations and potential corruption.
- The *institutional* heritage sector's capacity building can be assessed by paying attention to what educational courses in for example conservation, collection management, visitor's treatment and museum teaching are available or need to be developed in order to reach a high professional level in the sector.
- Evaluation has to adapt to Sierra Leone's fragile environment and pay attention to both its performance as a country and to donor organizations need for transparency.
- The *informal* heritage, or everyday heritage practices can be assessed by paying attention to everyday activities in designated areas/villages. In Sierra Leone for instance, one could observe local power structures – like Town Chief convening meetings, or cultural practices like Bondo in communities.
- Phenomenological approaches to people and landscape should be practiced in order to avoid pre-set participatory engagement, which generate a certain kind of knowledge, but not "deep" enough. The challenge is to convert this phenomenological knowledge into indicators and targets in evaluation for a systematic evaluation.

Values are quite central to all stakeholders, whether they are individuals, communities or institutions, and it has always been like that. In a fragile and conflict struck context, where (heritage driven) values are often targeted and destroyed, values can play a big role as building connections with stakeholders and may also be used as connectors rather than dividers. This is of course complex and diverse; values can unite internally and divide in relation to the outside. An evaluator needs to identify them by working with the people and their respective formal or non-formal institutions. Taliban, for example was a non-formal institution that existed in people's minds but later converted into a formal institution that was then formed to impose its values in communities' minds. Sierra Leone, on the contrary, has its heritage driven values that are partly composed by the colonial time values, to some extent through formal institutions that later became informal. Whether the local heritage driven values in Sierra Leone have anything to do with the years of conflict; or the formal institutions like Revolutionary United Front (that used child soldiers and hallucinogenic drugs for training purposes) played with local values and generated from a non-formal institution based on heritage-driven values are questions that need to be explored. The biggest questions for an evaluator in a fragile context are then how to tap into the heritage-driven values and how to identify if these values are dividers or connectors of conflict and if they are creating a non-formal institution in the communities. This is a highly delicate task as there is the risk of doing more harm than good, because, as argued in this essay, heritage is both omnipotent and double-edged. An evaluator, in such situations, is tasked to bring out local and national values driven from heritage perspectives and connect these to a sustainable peaceful future.

References

Appadurai, Arjun (ed.). 1986. *The social life of things. Commodities in a cultural perspective.* Cambridge: Cambridge University Press (Cambridge Studies in Social and Cultural Anthropology).

Aronsson, Inga-Lill. 2002. *Negotiating involuntary resettlement. A study of local bargaining during the construction of the Zimapán Dam.* Uppsala: Uppsala University, Department of Cultural Anthropology and Ethnology (Occasional papers 17). Doctoral thesis.

Aronsson, Inga-Lill. 2013. Det ambivalenta kulturarvet. Fallet terror Haza i Budapest. *Dragomanen* 2013 (15): 89-101.

Basu, Paul. 2008. Confronting the past? Negotiating a heritage of conflict in Sierra Leone. *Journal of Material Culture* 12 (2): 233-246.

Banton, Michael. 1969. *A West African city. A study of tribal life in Freetown.* London: Oxford University Press.

Bennett, Jon, Sara Pantuliano, Wendy Fenton et al. 2010. *Aiding the Peace: A Multi-donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan 2005-2010. Final report.* Hove, East Sussex: Itad. https://www.oecd.org/countries/southsudan/46895095.pdf

Cuno, James. 2008. *Who owns antiquity? Museums and the battle over our ancient heritage.* Princeton: Princeton University Press.

Lange, Matthew. 2015. States in the global south. Transformations, trends, and diversity. In: *The Oxford handbook of transformations of the state*, (eds) Stephan Leibfried, Evelyne Huber, Matthew Lange et al. Oxford: University Press (Oxford Handbooks Online). doi: 10.1093/oxfordhb/9780199691586.013.36

Okeke, Chukwuma O., Christopher N. Ibenwa, and Gloria Tochukwu Okeke. 2017. Conflicts Between African Traditional Religion and Christianity in Eastern Nigeria: The Igbo Example. *SAGE Open*. April-June 2017: 1-10. DOI: 10.1177/2158244017709322

Organisation for Economic Cooperation and Development (OECD). 2018. *States of fragility* 2018. Paris: OECD Publishing. https://doi.org/10/1787/9789264302075-en

Rodriguez-García, Rosalia. 2015. A portfolio approach to evaluation. Evaluations of community-based HIV/AIDS responses. In: *Poverty, Inequality, and Evaluation*, (eds) Ray C. Rist, Frederic P. Martin, and Ana Maria Fernandez, 139-174. Washington DC: World Bank Group. DOI-10-1596-978-1-4648-0703-9-EPI-210703.

Smith, Laurajane. 2006. *The uses of heritage*. New York: Routledge. https://rbb85.files.wordpress.com/2015/11/laurajane-smith-uses-of-heritage.pdf

Truth Commission Report (TRA). 2004. *Witness to truth. Report of the Sierra Leone Truth and Reconciliation Commission.* [Freetown:] Sierra Leone Truth and Reconciliation Commission. http://www.sierraleonetrc.org/index.php/view-the-final-report

Wilson, Richard A. 2001. *The politics of truth and reconciliation in South Africa. Legitimizing the post-apartheid state.* Cambridge: Cambridge University Press.

Evaluating transformational change — Lessons from international environmental funds

Juha I. Uitto, Jyotsna Puri, Anna Williams, Joe Dickman, Archi Rastogi, Geeta Batra and Kseniya Temnenko

ABSTRACT. Transformational change refers to deep, systemic, and sustainable changes with large-scale impacts in a significant area of concern, in the case of this chapter climate change and other global environmental issues. The chapter draws upon evaluative evidence from three major sources of finance for environmentally sustainable development, the Green Climate Fund (GCF), the Global Environment Facility (GEF), and the Climate Investment Funds (CIF).

To slow or stop further human-caused climate change and its devasting impacts on human and natural systems, and also to effectively prepare for it, there is wide recognition of the need for more transformational action than the incremental measures that have often dominated climate action to date. In line with their missions and underlying commitments, GCF, GEF and CIF have each been seeking to understand (positive) transformational change and how the work supported by each of these funds is supporting it. All three organizations are utilizing a framework for understanding transformational change that adopts (or in CIF's case, builds on) work done by the World Bank's Independent Evaluation Group in 2016 on transformational engagements, which focuses on four dimensions of transformational change that, when present together, signal that transformation is truly occurring. This paper explores each funder's perspective thus far on these concepts and analytical work done to date to assess whether climate interventions are thus far achieving transformation or have the potential or likelihood to do so.

Although the evaluative work undertaken by the three funders is in different stages and has taken different forms, there are clear indications that some investments have been – or are on a path to be – transformative, although not all investments are transformative and some are more advanced than others, and work to achieve transformation in very different ways. Many lessons as well as ongoing challenges are embedded in the wealth of knowledge and experience accumulated to date, and these are critical for informing more transformational ongoing and future investments.

Several challenges nonetheless are inherent in the evaluation of transformative change, including issues of methodology, scope, complexity, resources, and technical capacity. However, the opportunities for advancement are many, including collective efforts across funders and other practitioners, and the growing body of knowledge and strategic learning on transformation change in climate action.

Introduction

In 2015, the leaders of UN member States agreed on the comprehensive 2030 Sustainable Development Agenda and the accompanying Sustainable Development Goals (SDGs) that apply to all signatories, both developed and developing countries. In the same year nations came together around the Paris Agreement to limit global warming. Despite these political commitments, global environmental trends continue to point downwards. Biodiversity loss has reached catastrophic levels with significant and irreversible consequences to ecosystem integrity and functions (Caballos, Ehrlich and Dirzo 2017). The impacts of climate change are already visible in terms of unprecedented global temperatures, weather anomalies and extreme events, melting ice sheets, and rising sea levels. The Intergovernmental Panel on Climate Change (IPCC) warns that controlling climate change will require rapid and far-reaching transitions in energy, land, urban and infrastructure (including transport and buildings), and industrial systems, not to mention disproportionate effects on vulnerable people and high rates of species extinction (IPCC 2018).

Given these dire scenarios, business as usual will not be sufficient. Consequently, many organizations working on global environmental and development issues have set their sights on promoting transformational change. The goal of reconciling the demand for economic growth and shared prosperity, while protecting the environment and maintaining the natural resource base, is of course not new. The first UN conference on the Human Environment was held in 1972, spurred on by concerns raised in reports by the Club of Rome and others (Meadows et al. 1972). The Brundtland Commission released its report, *Our Common Future*, in 1987, introducing the term sustainable development (WCED 1987). The UN Conference on Environment and Development, known as the Earth Summit, in 1992 brought together an unprecedented number of heads of state and government, and led to the establishment of Rio Conventions, including the Convention on Biological Diversity (CBD) and the UN Framework Convention on Climate Change (UNFCCC) to which the Global Environment Facility (GEF), founded in 1992, became the financial mechanism.

The Climate Investment Funds (CIF) were established in 2008 to scale up finance for climate change mitigation and resilience, filling urgent financing gaps and demonstrating the viability of emerging solutions. With more than US$8 billion contributed, CIF's goal is to advance transformational change toward low-carbon, climate-resilient development. The GCF was established by the Conference of Parties in 2012 as an operating entity of the financial mechanism of the UNFCCC,

and one its objectives is "to promote a paradigm shift towards low-emission and climate-resilient development pathways".[1]

The emphasis on transformative change is necessarily ambitious. Achieving positive transformation in climate action will require ongoing commitment and novel strategies and approaches. While the scientific foundation is rather clear, evaluation can shed light on the types of policies, programs, and interventions that have been transformational or show promise in this regard. Learning from the past and present interventions is critical to enhance the transformative impact of environment and development efforts.

In this chapter, we draw upon work conducted under the auspices of GEF, CIF, and GCF, three major international funds aiming to support transformational change in least-developed countries and economies in transition. The first section of the paper presents a conceptual framework of how to define and measure transformational change, drawing on earlier work by the World Bank Independent Evaluation Group (IEG 2016) and related emerging work at GCF and elsewhere. The following sections present results from the Evaluation of GEF Support to Transformational Change and findings from the CIF Transformational Change Learning Partnership (TCLP). We conclude by describing some common challenges, lessons, and directions for the future.

Paradigm shift promoted proactively – the case of the GCF

Paradigm shifts have occurred throughout human history. For instance, Lewis and Maslin (2015) identify critical revolutions, including the rise of agriculture (11,000-5,000 BC) and industrialization (eighteenth century to today), with years 1604 or 1984 proposed to mark the beginning of the Anthropocene Epoch. While many paradigm shifts have occurred through human history, one of the key sources of the concept of a transformational change or paradigm shift is Thomas Kuhn's classic work on scientific revolutions (Kuhn 1962), where transformation in scientific inquiry is viewed to occur in three phases. In the first 'pre-paradigmatic phase,' scientific activities and most conceptual development take place within the pre-existing paradigm. With time, this leads to the second phase of 'normal science,' with the development of experimentation and data in the dominant paradigm, aided by new scientific techniques and technology. In

1 In the GCF discourse, the term 'paradigm shift' is used instead of transformational change. These terms are used interchangeably in the rest of the chapter.

this watershed phase, ongoing scientific inquiry may confirm the dominant paradigm.

On the other hand, anomalies may emerge as scientific inquiry develops, and accumulation of such anomalies may lead to a revolution, a transformation or a paradigm shift. Outside of scientific inquiry, Kania et al. (2018) argue that systems change requires six conditions: policies, practices, resource flows, relationships/connections, power dynamics, and mental models. They suggest that only the first three are explicit, while shifts in mental models are implicit (relationships/connections and power dynamics are regarded as semi-explicit). The authors further suggest that while it is necessary to work at all levels of change, measurable change at the tip of the proverbial iceberg of transformation will become evident later in the process. Another implication of this model is that it is important to build the explicit, while the implicit transformative changes become evident later in the process. This is not to say that this model of change directly applies to the GCF. However, it is too early to see evidence of measurable change as a result of GCF activities. Much of this change is implicit as GCF activities have focused on building an institutional infrastructure and a pipeline of projects. Other authors have described transformational changes through the lenses of transition theory and resilience theory (Ferguson, Brown, and Deletic 2013) and social ecological systems (Moore et al. 2014).

In the next section, we focus instead on insights on evaluating transformational change from the perspective of a young organization.[2]

Transformational change in the GCF discourse

Among the organizations considered in this chapter, GCF is the most recent to be established. The GCF was established by the Conference of Parties in 2012 as an operating entity of the financial mechanism of the UNFCCC, and one of its objectives is "to promote a paradigm shift towards low-emission and climate-resilient development pathways". The Independent Evaluation Unit (IEU) of the GCF commenced its evaluation of the paradigm shift with a review of instances where an attempt was made to define and measure transformational change (Puri 2018). Through a review of multilateral and bilateral agencies with a stated focus on transformational change, the IEU found some common features in the experience

2 This section draws from evaluative studies undertaken by the IEU. These include: a learning paper to examine how and if transformational change can be defined and measured (Puri 2018), and a synthesis study underway to collate, critically appraise and synthesize available documented evidence and recommend key areas for further examination. The latter is part of the overall Performance Review of the GCF, requested by the GCF Board and being undertaken by the IEU in 2019.

identifying transformational change in various agencies (Table 1). Of note, behaviour change or 'the last mile problem' is not given sufficient consideration in this exposition, in our view. Indeed, the Great Transformation was the topic of the first Nobel Laureate Symposium and for the Potsdam Memorandum, where Gell-Mann (2010) identified behaviour change as one of the fundamental factors for transition towards sustainability.

TABLE 1. A review of experience in identifying "transformational change" across evaluations

Attribute of T-change	CIF trans-formational change studies	WB trans-formational engagement	GEF LDCF/ SCCF	UKCIP	IFAD
Measured T-change?	No	Maybe	No	No	Yes
Specific/consistent indicators	No	No	No	No	Yes
Demonstration project logic (toc)/ catalytic	Yes	No	Yes	Yes	No
Removing barriers/ lower costs	Yes	No	Yes	?	No
Scale effects (spatial)	Yes	Yes	Yes	?	Yes
Research and learning	Yes	No	Yes	?	No
Systems and across sectors	Yes	Yes	Yes	Yes	Yes
Long-term change	Yes	Yes	Yes	Yes	Yes
Behaviour change	Yes	Yes	No	No	Yes
Capacity building	No	No	Yes	No	No

SOURCE: PURI 2018

As stated previously, the GCF is proactively mandated to promote a paradigm shift. The Governing Instrument of the GCF identifies that one of the objectives of the GCF is to promote a paradigm shift towards low emission and climate-resilient development pathways. During the process of its institutional development, the GCF has operationalized this paradigm shift at many levels. The Initial Strategic Plan of the GCF recognizes the challenge to turn this "abstract vision into practice." The plan further accords the promotion of a paradigm shift as one of two Strategic Vision statements, with some components of a paradigm shift highlighted. Further, the GCF has articulated an Initial Investment Framework, which provides guidance to developers of Funding Proposal. This framework includes paradigm shift as one of the six investment criteria against which proposed GCF investments are assessed. This framework defines paradigm shift potential of a funding proposal as: "degree to which the proposed activity can catalyse impact beyond a one-off project or program investment." At the time of writing this chapter, the GCF is undertaking planning for the first strategic period, as well as a review of the investment framework. A detailed performance review led by the IEU was undertaken in 2019, which reviewed the likelihood of impact of GCF investments.[3] Based on a desk-based study, this chapter focuses on insights, which are generalized to multilateral organizations seeking to promote a paradigm shift.

Although a relatively new agency, the GCF has built a strong portfolio, committing US$5.0 billion to 102 projects focusing on eight result areas, spanning adaptation and mitigation. This portfolio of projects addresses the multifaceted mandate of the GCF. While the primary objective of the GCF is to promote a paradigm shift, it is also mandated to pursue this through direct access, while unlocking the potential of the private sector, in a way that responds to country needs, accounting for gender and environmental and social safeguards, building capacities in developing countries, and through a balanced portfolio. An initial document review identifies 19 priorities within the GCF architecture. Arguably, different GCF projects would emphasize different sets of these priorities. As a critical mass of projects builds over time, it is important to recall two associated factors:

While the mandate of multilateral organizations can be multifaceted, each individual project may contribute only to part of the mandate. To illustrate, it is possible that a GCF project may do well on one investment criterion, and less well on another criteria. As the mass of projects develops, it can be expected that the collective impact of projects will catalyse a paradigm shift in multiple investment criteria.

3 See Performance Review of the GCF: https://ieu.greenclimate.fund/evaluations/fpr

For multilateral organizations with multifaceted mandates, it may be important to consider the investment criteria and accord them levels of priority.

Related to the above point, it is possible that there may be parts of a mandate which align with each other. The IPCC in its latest analysis suggests the pursuit of adaptation goals is consistent with those of the SDGs, and that there are synergies between the strategies to achieve both. The transformational changes supported by the achievement of the SDGs would also be necessary for limiting warming to 1.5°C above pre-industrial levels (Roy et al. 2018). Therefore, for a young organization there may be benefits from focusing on areas where multiple or synergistic benefits can be accrued. This aligns also with the thinking where transformational change is viewed as polycentric in nature.

Evidence shows the need for clear guidance, and strong measurement tools while building for a paradigm shift. The GCF is specifically mandated to carry out its activities through direct access and responding to the needs of countries. In operational terms, this means that proposals are brought forward by Accredited Entities for review by the GCF Board. Currently, there is not enough evidence to suggest whether or not this ensures that the resulting proposals are aligned with the GCF mandate, or whether or not the vision of the proposing entity aligns with that of the GCF. On a related note, a working paper of the IEU finds that between 16-39% of funding proposals rely on significant assumptions that are not verified and/or where paradigm shift indicators are vaguely described (Fiala, Puri and Mwandri 2019). This has two implications. Firstly, this points to the need for clarity in guidance and exposition on a paradigm shift. This is especially necessary if the scope of the organization is vast (or indeed global), and its activities are to be carried out through a large variety of entities with varied organizational understanding of paradigm shift. Clear guidance on what the organization seeks will lead to a congruent pipeline. Secondly, it is also necessary for such organizations to build measurement frameworks, so that credible evidence is systematically collected for accountability and learning. Coming back to the "iceberg of transformational change", while the explicit is being built, it is important to create measurement frameworks that can credibly and sufficiently measure the implicit change, as and when it may occur.

There is evidence in the literature to suggest that a paradigm shift occurs under one of the two scenarios: (a) when it is not driven proactively but catalysed by a combination of factors, in examples such as the Neolithic Revolution (the diffusion of arable farming and animal husbandry) and the Industrial Revolution (the transition from an agrarian to an industrial society), or (b) is actively driven by entrepreneurs,

like the examples set by Microsoft or the IT sector. In neither example is a paradigm shift achieved through a mandate given to a specific body, and instead could be viewed as the 'uncontrolled results of evolutionary change' (Schellnhuber et al. 2011). How, then, can a paradigm shift be proactively pursued? Schellnhuber et al. (2011) recommend, among other things, improving proactive states with extended participation opportunities. Within the GCF architecture, this is provided for through enhanced country-ownership. Country ownership, although not specifically defined, is indeed seen as an important component of a paradigm shift.

Finally, we also identify four features of a paradigm shift, that are necessary but not sufficient, as an organization develops its institutional infrastructure (Table 2). We draw these from the work undertaken by the World Bank (2016) (see also the section on the GEF below).

TABLE 2. Dimensions of transformational change, Independent Evaluation Group, World Bank 2016

DIMENSION	DESCRIPTION	CHARACTERISTICS
Relevance	Addresses a major development challenge (or societal or global concern) such as poverty, equity.	Evidence from diagnostic or analytical work showing the constraint or problem addressed was of critical importance.
Depth of change	Causes or supports fundamental change in a system or market; addresses root causes to support a change in trajectory.	Evidence of market change, systemic change or behavioural change.
Scale of change	Causes large-scale impact at the national or global level.	Evidence of scaling up of approaches and innovations and replication; catalytic effects; demonstration effects; positive spill-overs and externalities; acceleration/discontinuity in a development indicator.
Sustainability	Impact has been economically, financially, environmentally sustainable in the long run.	Evidence of financial, economic, environmental sustainability of results after engagement.

SOURCE: AS STATED IN PURI 2018

On the basis of the above dimensions of transformational change, we propose that transformational change is possible when it is relevant – it addresses a major global concern such as climate change. Organizations focusing on transformational change would do well to focus on depth of change, such that the intervention addresses a root cause and creates a fundamental change in the trajectory. We further suggest building for scale, and focusing on large scale impacts at the national or global level. Finally, in promoting a transformational change, organizations may focus on the sustainability of initiatives so that the impact of such change is viable for the long-term.

Evaluating GEF support to transformational change

The oldest of the three financial mechanisms covered in this chapter, the GEF has supported more than 4,500 projects and provided over US$17.9 billion in grants and mobilized an additional US$93.2 billion in co-financing in the areas of biodiversity, climate change, land degradation, international waters, and chemicals and waste since its establishment in 1991. Promoting transformational change is a strategic priority of the GEF, as cited in its 2020 vision statement (GEF 2015). This focus builds upon a quarter century of experience with programming. Evaluative work by the GEF Independent Evaluation Office (IEO) conducted in the context of the Sixth GEF Overall Performance Study (IEO 2014) articulated a general theory of change for the GEF that identified the outcomes to impact pathways. In this regard, the broader adoption of actions and the behavioural change initiated by GEF-funded projects would be the transformational processes leading to impact (IEO 2013).

In 2017 IEO conducted an evaluation to review GEF experiences in promoting transformational change and to identify contributing factors (IEO 2017).[4] The four specific criteria outlined in Table 2 above were used to differentiate transformational interventions from those that are "merely" highly successful, complex or large in size: (i) relevance; (ii) depth of change; (iii) scale of change; and (iv) sustainability. The underlying theory of change applied was that by strategically identifying and selecting projects that address environmental challenges of global concern and are specifically designed to support fundamental changes in key systems or markets, GEF interventions are more likely to cause a sustainable large-scale impact, assuming supportive contextual conditions, as well as good implementation of the project (Fig. 1).

The specific evaluation questions were:

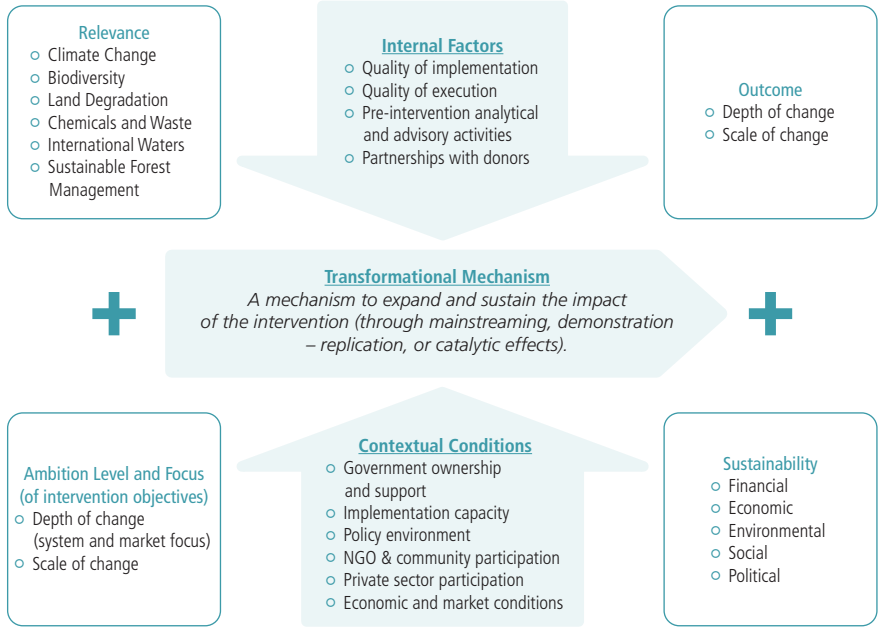
- What are the necessary and sufficient conditions for GEF interventions to achieve transformational change?
- What causal factors make a difference in the outcomes?

The evaluation was based on a purposive sample of projects that had been selected from a set of 156 completed projects nominated by GEF Agencies[5] at the request

4 The core authors of this evaluation were Andres Liebenthal, Geeta Batra and Kseniya Temnenko.

5 The World Bank, United Nations Development Programme (UNDP), UN Environment, the Food and Agriculture Organization (FAO) of the UN, and the Asian Development Bank (ADB).

FIGURE 1. Theory of change for GEF transformational interventions



SOURCE: GEF IEO 2017

of IEO. These were screened for meeting the criteria for transformational change – defined as deep, systemic, and sustainable change with large-scale impact in an area of global environmental concern – verified through independent project-level evaluations. Eight illustrative projects were selected taking into account their diversity in focus, regional distribution and agency:

- Lighting Africa
- China Renewable Energy Scale-Up Program, Phase I (CRESP-I)
- Uruguay Wind Energy Program
- Sanjian Plain Wetlands Protection Project (China)
- Sustainable Land, Water, and Biodiversity Conservation and Management for Improved Livelihoods in Uttarakhand Wind Sector Project (India)
- Strengthening the Projected Areas Systems in Namibia

- Amazon Regional Protected Areas Program, Phase I (ARPA-I) (Brazil)
- Promoting Payments for Environmental Services and Related Sustainable Financing Schemes in the Danube Basin

The evaluation used a cross-case analysis in combination with a meta-analysis of the project-level evaluations for each of the projects to assess the conditions, and combinations of conditions, that support transformational change.

Based on the analysis, the IEO was able to identify a number of factors common to these projects that provide lessons for the future. First, these interventions that achieve transformational change had clear *ambition at design*. Their objectives were set to aim at a profound, fundamental change in addressing a market distortion or a systemic bottleneck that was a root cause for a global environmental concern. Four of the cases focused primarily on systemwide transformation, taking a comprehensive approach to modifying the functioning of a collection of components (economy, public sector, private sector, community) that interact with one another with environmental consequences. In the four other cases, the primary thrust was on transforming a market (i.e., the supply and demand of goods and services) associated with environmental impacts of global concern.

Secondly, the adequacy of the policy environment had an important impact on the depth and scale of reforms promoted by all transformational interventions. All projects thus addressed *market and system changes through policies*. In three cases, the interventions had a major role in helping define and implement the main policies essential to trigger and sustain transformational change. In China and Uruguay, the projects had a strong influence on policies that provided an effective stimulus to the development of renewable energy. In Namibia, the project provided technical support for the drafting of new policies for the Ministry of Environment and Tourism, affecting the protected area systems. In three other cases, the interventions played a modest role in strengthening the policy framework needed to support transformational change. These ranged from discussing with the governments of Ghana and Kenya lowering impact taxes as an enabling environment for the solar lamps market, to proposing prohibition of animal grazing and fishing in all nature reserves in the Sanjiang Plain in China, to mainstreaming payment for environmental services concepts into national fisheries policies in Bulgaria and Romania. In India, the state government of Uttarakhand granted the local rural governments formal legal recognition for watershed development planning. In Brazil, ARPA-I used the existing legal context for protected areas to involve many government agencies and financing partners

to demonstrate the practicality of a participatory approach to the establishment and management of protected areas.

Thirdly, the interventions had established a *mechanism for financial sustainability* by integrating within government budgetary systems or by leveraging market forces and key stakeholders' economic interests. In China, CRESP-I supported a feed-in tariff for renewable energies that provided financial returns attractive enough to encourage state-owned and private enterprises to accelerate their investment in renewable energy (Box 1). In Uruguay, wind power investment licenses were allocated through a fair bidding process that guaranteed access to the grid. The resulting prices were competitive with those of fossil fuel alternatives and have gradually declined further as a result of growing efficiencies and technological improvements. In Uttarakhand, project beneficiaries have an incentive to maintain water harvesting structures, namely their own investment through cost-sharing. In the Sanjiang Plain, a portion of local county revenues generated from forest development activities are used to meet the financing requirement for nature reserve management.

Fourthly, another common feature was that all interventions were well implemented in terms of quality of project design, supervision by the GEF Agency, and the effectiveness of the executing agencies. Some of the salient features that had driven the quality included: comprehensive diagnostic assessments to identify barriers to be addressed; coherent designs to target all identified barriers; early involvement of strong executing agencies that would own the project objectives and be willing to learn, adjust and adapt the design, scope and management as needed to ensure success.

Interestingly, transformation can be achieved by projects of different sizes. While most of the projects included in the sample were multimillion-dollar efforts with long durations, the projects in Uruguay and the Danube Basin were relatively small with a limited duration but targeting important barriers and working with key stakeholders at the right time.

The purpose of this evaluation was primarily for learning. The evaluation framework and findings can help the GEF and potentially other organizations to gauge project concepts in advance to determine their probability for supporting transformative change and how project designs can be enhanced with this in mind.

BOX 1. Scaling up China's renewable energy sector

China's energy consumption and associated carbon emissions had been rapidly increasing in the decades before the launching of the CRESP-1 project in 2005. The World Bank and the GEF worked closely with the Chinese government to develop a long-term partnership to increase the contribution of renewable energy to power generation in a sustainable way. The project was designed as a programmatic and sector-wide intervention that integrated a GEF grant (US$40.2 million) to support the development of the legal, regulatory and policy framework needed to stimulate demand for renewable energy, improve its quality, and reduce its costs, and to build a strong local renewable energy equipment manufacturing industry, with two World Bank loans (US$87.0 million and US$86.3 million) to support pilot investments in wind, biomass and small hydropower in four participating provinces. These ambitious objectives aimed at major changes in the system and market for renewable energy in China. Five years after the project closing the World Bank Independent Evaluation Group (IEG 2017) concluded that CRESP-I had made a substantial contribution to the transformation of China's renewable energy sector from an early piloting and demonstration stage to the country's status as a global leader in wind energy generation and the manufacture of wind power equipment. From 2005 to 2010, China's installed wind power capacity increased from 1.3 GW to 29.6 GW, and further to 129.3 GW in 2015, amounting to 22% of global wind power capacity and 3% of China's electric power generation. These impacts are likely to be sustained given the government's implementation of project-recommended tariff policy and its commitment to further increase the share of non-fossil fuels to 15% by 2020.

Key stakeholders consulted credited the project's instrumental role in tariff-related studies, which provided the knowledge and analytical underpinnings for the replacement of a project-by-project tariff-setting and concession system with a national tariff structure that offered attractive and predictable returns to investment. The evaluation concluded that the main factors that contributed to the project's transformational impact included: (i) the integration of institutional development and capacity building, technology improvement, and investment activities in a single intervention with mutually reinforcing components; (ii) extensive efforts through workshops, study tours and studies during a multi-year preparation period to achieve consensus and cohesiveness about key policy directions and reforms; (iii) cost-shared sub-grants (where the grant provided 20-25% of total research and development costs) that leveraged substantially greater investments by the counterparts, enhanced selectivity, and build ownership and commitment; and (iv) the long-term, predictable and financially attractive price signal implemented by the government, as recommended by project-supported studies.

CIF Transformational Change Learning Partnership

The Climate Investment Funds (CIF) were established to scale up finance for climate change mitigation and resilience, filling urgent financing gaps and demonstrating the viability of emerging solutions. With more than US$8 billion contributed since 2008, CIF supports transformational change toward low-carbon, climate-resilient development in the areas of mitigation, resilience, and forests through four programs: The Clean Technology Fund (CTF), the Pilot Program for Climate Resilience (PPCR), the Forest Investment Program (FIP), and the Scaling Up Renewable Energy in Low-Income Countries Program (SREP).[6] These programs have supported 300 projects across 72 countries.

CIF established the Transformational Change Learning Partnership (TCLP)[7] in 2017 to facilitate a collaborative, evidence-based learning process on transformational change in the CIF context. The TCLP concluded its current phase work in May 2019, and at the time of this writing, was exploring options for extending some elements. The reflections below cover the TCLP work between the spring of 2017 and the spring of 2019.

The TCLP had three related components:

1. An independent evaluation of transformational change;
2. An independent evidence synthesis of transformational change; and
3. A facilitated learning process with CIF stakeholders.

All TCLP components aimed to understand CIF's role in contributing to transformational change by answering four questions on concepts, process and design, results, and learning:

(1) ***Definitions*** How is transformational change conceptualized in the international field of climate finance?

(2) ***Process and design*** To what extent and how does CIF's approach to planning, designing, and implementing its investments work to advance

6 See for an overview and background of these programmes https://www.climateinvestmentfunds.org/.

7 TCLP (https://tinyurl.com/y29fmvbb) is part of the CIF's Evaluation & Learning Initiative. (https://tinyurl.com/y5d6m24h).

transformational change?

(3) ***Results*** To what extent, how, and under what conditions are CIF-supported investments and activities contributing to transformational change?

(4) ***Learning*** How can CIF and others increase their contributions to transformational change?

A first task was to identify a working definition of transformational change, building on previous work by CIF (2015) and others, including the World Bank Group's Independent Evaluation Group, and the GEF's Independent Evaluation Office, which had both recently assessed the transformation influence of project and programme interventions. The TCLP defines transformational change in climate *action as strategic changes in targeted markets and other systems, with large-scale, sustainable impacts that shift and/or accelerate the trajectory toward low-carbon and climate-resilient development*. During the first year of the TCLP, the World Bank's initial definition of transformational change was expanded upon to identify four similar dimensions of transformational change (relevance, systemic change, scale, and sustainability; see Box 2) that must all be achieved to realize comprehensive transformation.

The TCLP's interpretation modestly varied from or perhaps refined the four dimensions, at least to be well suited to the TCLP's purposes. For example:

BOX 2. Four Dimensions of Transformational Change

RELEVANCE The strategic focus of CIF investments: impacting low-carbon and climate-resilient development, with sustainable development co-benefits.

SYSTEMIC CHANGE Fundamental shifts in system structures and functions.

SCALE Contextually large-scale transformational processes and impacts.

SUSTAINABILITY The robustness and resilience of changes.

Relevance (TCLP)/Relevance (WB IEG): The TCLP's interpretation goes further in ambition, such that an effort not only addresses a major challenge, but that it must also be strategically focused on removing barriers and advancing opportunities for transformation. (Thus, for instance, not only related to climate change, but strategically relevant to unlocking transformation, beyond business-as-usual or incremental approaches.) The TCLP's definition of relevance also adopted language from CIF's original mission statement and recognizes broader human, economic, and environmental "co-benefits" as integral to transformative relevance.

Systemic change (TCLP)/Depth of change (WB IEG): The two articulations use somewhat different terms but are similar in emphasis.

Scale of change (TCLP)/Scale of change (WB IEG): Definitions are similar, both emphasizing large-scale impacts. TCLP definition does not specify national or global specifically, in recognition that large scale can take a number of forms, such as technology, sector, or regional, even if ultimately global is the ultimate goal.

Sustainability (TCLP)/Scale of change (WB IEG): The TCLP definition clarifies that "sustainability" is not necessarily clear as a term on its own, and that in this context is about robustness and resilience to future setbacks and barriers, such as discontinuation of any concessional financing, shifts in markets, unpredictable disasters, etc.

The TCLP's version of these definitions are expanded upon in supplementary materials that explore both impacts and processes for each dimension. While the TCLP work on these dimensions arguably advanced the prior IEG work, at least for the CIF TCLP context, there is still room to continue advancing these definitions support, for example, to more explicitly address the speed of change, which is not directly addressed by the current four dimensions. There are also other definitions and frameworks that are emerging as well as as-of-yet untapped research and work from other sectors that could further inform more work on transformational change in climate action – and beyond, given the relevance to practically any complex systems change, regardless of sector or context.

TCLP Studies and Facilitated Learning Process

The evaluation of transformational change in the CIF analysed CIF's contribution to transformational change by assessing CIF's work in a purposeful sample of countries and programs against the four dimensions of transformational change, as well as testing hypotheses across case studies in 15 countries,[8] supplemented by information from other countries and sources, including interviews with over 250 individuals. The evidence synthesis focused on systematic collection and screening of secondary literature, considering factors including the four dimensions as well

8 The evaluation findings focus on CTF programs in Chile, Mexico, Morocco, Thailand, and Turkey; SREP programs in Armenia, Honduras, Kenya, and Nepal; PPCR programs in Jamaica, Mozambique, Nepal, Niger, Tajikistan, and Zambia; and FIP programs in Burkina Faso, Mexico, and Mozambique.

as CIF design and other factors. The synthesis extracted evidence related to CIF's role in transformational change from over 85 sources. Although the two studies differed in methods and information sources, the findings are broadly aligned, and are therefore combined for this summary. The facilitated learning process brought together over 60 CIF and external stakeholders over the course of four workshops and several online events to help inform the conceptual framework to validate early findings and deepen ongoing learning and utilization outcomes.

Contextualizing the findings

Program differences: It is important to consider differences in CIF program funding levels, sectors, design, and implementation status when interpreting the findings. The CTF portfolio is the largest and most advanced, with 70% of projects approved between 2009 and 2015. The PPCR portfolio is maturing, with 60% of projects approved between 2013 and 2015, but is largely still at an early stage of implementation. Over half of SREP and FIP projects are only one to two years old, with some generating results but most still at the design or early implementation stages.

Limitations: There are inherent limitations to analyses of complex systems change of this magnitude, including isolating CIF's influence, generalizing findings based on case studies, and potential bias stemming from a focus on more advanced programs. The evaluation team limited potential bias by gathering input from external experts; identifying positive, neutral, and negative dynamics; triangulating evidence; and caveating findings where appropriate. The evidence synthesis drew on publicly available literature[9] on CIF's role related to transformational change.

The world has changed since CIF's founding: For instance, in 2008, low-carbon energy technology costs were high, penetration of these technologies in low- and middle-income countries was limited, and the enabling environment was not conducive to scaling. In the decade since, there have been large-scale investments in renewable energy technologies – notably wind and solar PV – with plummeting technology costs and increasing penetration in emerging markets.

9 The evidence synthesis drew from publicly available literature published in English, which is inherently limited. It did not incorporate internal MDB project documents, which was outside of the scope. These documents generally do not examine systems-level transformational change as defined by the TCLP.

The concept of transformational change has also evolved. The current TCLP concepts were retrospectively applied to work initiated ten years ago, when the four dimensions of transformational change were not identified. External factors also affect progress, including wider political, social, and environmental events, with investments occurring in complex and often uncontrollable contexts. Progress is often non-linear and unpredictable.

Overarching Findings on CIF's Contributions to Transformational Outcomes

There is evidence of progress toward transformational change across CIF programs to greater and lesser extents, depending on the program and dimension of change. The evaluation classified this progress in terms of maturity in the signals of transformation (see Table 3). Advanced signals indicate strong evidence of transformation, interim signals indicate transformation is underway, and early signals indicate pre-conditions are in place for future progress.

TABLE 3. Signals of transformational change by dimension and program

Dimension of transformational change	STAGE OF TRANSFORMATIONAL SIGNALS			
	CTF	SREP	PPCR	FIP
Relevance	Advanced	Advanced	Advanced	Advanced
Systemic change	Advanced	Interim	Advanced	Interim
Scaling	Advanced	Early	Interim	Early
Sustainability	Advanced	Early	Interim	Early

This table reflects the balance of evidence collected through the evaluation across the countries covered. It is not an assessment of the overall portfolio.

SOURCE: MCPHERSON 2019

The evaluation and evidence synthesis both found that CIF's country programs are generally well designed and highly relevant to supporting transformation. This is also reflective of GEF's findings that strong program design can enhance the chances of transformational outcomes. The studies also found that other specific elements of the CIF design were particularly conducive to supporting transformation, including: a country-led programmatic approach (CIF and ICF 2018); explicit consideration of transformational change at the design phase; large-scale investments utilizing a range of concessional financing tools; delivery of financing through multiple,

coordinated multilateral development banks (MDBs); and a flexible and predictable funding envelope.

At the program level, the studies found that CTF has realized the most transformative results thus far. Operating in more-mature markets (Mexico, Turkey, Thailand and Morocco), CTF country-level programs commonly demonstrate advanced signals across all four dimensions, where low-carbon energy has shifted toward non-concessional, market-based approaches. The strong signals of scaling and sustainability reflect the development of private investment and developer markets.

The Strategic Climate Fund programs (PPCR, SREP, and FIP) have made interim or advanced progress on systemic change, reflecting changes in institutional structures and functions, even behaviour change. For PPCR, fundamental shifts in stakeholder behaviours, knowledge, and capacity demonstrate advanced systemic change, in for example Jamaica, Mozambique, Niger, Tajikistan, and Zambia. PPCR also commonly sees interim signals of scaling and sustainability, reflecting the mainstreaming of climate change into government structures, decision making, and budgeting, although there is variation in progress between countries, with the transition to sustainable models occurring at different speeds.

Differences in progress between programs hinge upon a range of factors, including implicit differences in sectors and program design, complexity, and level of resources available. Early and interim signals were more common in less-developed country contexts, where capacity and governance can represent constraints, and in more complex and contested thematic areas with strong socio-economic linkages (e.g., forests, community resilience, low-carbon transportation). It is generally also easier to measure advanced signals of change on technology deployment, investment, and finance (e.g., CTF), than through indicators of resilience or measures of systems change.

Anchoring CIF programming in a narrative of wider co-benefits has helped support transformational change in local contexts. Such benefits include reducing poverty, a key driver for many low-income country governments—particularly concerning community adaptation, forest livelihoods, and energy access agendas. They also include economic development and industrial green-growth strategies, including manufacturing, job creation, and higher productivity. Some investments do not (at least as of yet) show signals of transformational change, particularly in the less mature FIP and SREP portfolios. This is in part because it generally appears to take four to seven years from project approval for implementation to be underway long enough for transformation to have the chance to emerge. Country-level barriers to transformation—such as institutional capacity constraints, subsidies for non-

sustainable alternative agendas, political instability, catastrophic weather events, and insufficient long-term financing—also inhibit progress.

Reflections on transformational change processes and timelines

The evaluation included reflections on how transformational change occurs when a range of context-specific factors align and can vary significantly by program, theme, or market. As such, it can appear chaotic and unpredictable, particularly when comparing interventions or timescales across a broad portfolio.

Incremental change from individual projects should not be outright discounted, as it can also cumulatively make future transformational tipping points more likely; however, realistic expectations should be applied regarding both the timescales for transformation and the role that specific interventions can play, especially in more complex areas or contexts.

The evidence synthesis revealed that there is an overarching lack of sufficient research and analysis on the role of CIF (and other climate finance institutions) on transformational change, though recent work commissioned through the CIF Evaluation and Learning Initiative and undertaken by partner institutions including GCF, GEF, and others is beginning to address some of these gaps. Still there are remaining knowledge and evidence gaps, some of them large (e.g., forest and land use, including coverage of the FIP program in general), private sector investments in resilience, transport, and others.

Both studies recognized that achieving transformation requires a commitment for the long-term, including commitment to strategic yet experimental and sometimes risky investments that have the potential to remove barriers and push the frontiers of progress.

Recommendations to CIF and other climate finance stakeholders

The evaluation and evidence synthesis concluded with recommendations in support of strengthening the transformational potential of climate finance. The recommendations include developing tools to support transformative programming design (e.g., such as guidelines, frameworks, and country-level theories of change); approaching transformation from a portfolio perspective; and supporting investments that address more complex barriers and require longer-term support, thereby maximizing the benefits from flexible, concessional finance to support priority challenging and emerging areas.

Overarching reflections on transformational change

The CIF transformational change evaluation and evidence synthesis were ambitious efforts undertaken within a short timeframe, utilizing a limited set of secondary data and also limited primary data collection given the scope of the study and the CIF portfolio. That said, both efforts, despite their limitations, revealed important patterns around transformational trends and tipping points, as well as significant differences in how transformation occurs in different programs, sectors, and contexts. Both studies, and the broader community working on this, recognize that although the four dimensions of transformational change serve as a strong basis to understand and assess transformation, the concept of sustainability remains both the most elusive and the hardest to maintain or guarantee over time, particularly when the initial concessional investments were completed years prior.

The broader TCLP work has been viewed by participants as both engaging and useful for building a shared understanding of this complex topic as well as illuminating lessons on how transformational change has (and has not) occurred and CIF's role in this change. The TCLP is now reflecting on the work over the past two years and exploring options for future activities. There are many remaining questions, outstanding knowledge gaps, and a time-sensitive global mandate to ensure that climate finance is most effectively used to maximize positive transformation. Given the urgency of the climate crisis and unprecedented investments now being made to prevent and prepare for it, the time to continue deep learning on this topic and improve the transformational impacts of climate action is clearly now.

Conclusions

The experiences on evaluating transformational change from the three major sources of climate and environmental funding point to the potential such financing can have in moving systems and markets towards more sound and sustainable development trajectories. We have identified common factors that define transformational projects and programs. These include targeting critical issues at the systemic level from the outset with the goal of achieving depth and scale of change. It is also important to focus on strong program and project design, with explicit considerations for sustainability.

There are several challenges surrounding evaluation of transformational change, starting with the nature of the climate crisis, the epitome of a complex

"wicked problem" with many interdependent factors, largely uncontrollable and unpredictable dynamics, and no feasible solutions thus far. Thus, the nature of the problem is a primary challenge.

The next challenges surround the sheer depth and breadth needed to solve the crisis. To be effective in the long run, interventions intended to mitigate and prepare for climate change need to work at all levels of society, to affect the kinds of real systems change needed. But even then, both climate change and human society are complex dynamic systems, effectively layering one complexity challenge upon another.

Evaluating the effectiveness of such systems-change interventions is thus inherently difficult, and it must go beyond both the direct effect of investments on a discrete, defined target population (where this is even possible given that many interventions are focused on enabling environments, capacity building, policy advancements, etc.). They must examine how broader sectors and society – beyond direct investments – are improved or influenced to support positive and transformative systems change.

Further, to be fully realized, transformational change must also occur well past the lifespan of any climate finance investment, and in fact become self-sustaining by broader society. Therefore, any advanced evaluation of transformational change must also examine these realities by looking at systems change beyond individual investments in terms of scope, scale, and time. This is not what evaluations typically do, nor do we have the mandate or types of evaluative tools for this.

This is not only a methodological challenge. The challenges are also about budgets, data availability, and skillsets. Undertaking systems change evaluation on climate change investments at a global scale is very ambitious. Budgets needed to do this would ideally be commensurate with the challenge, but these resources are all but readily available. The data (both quantitative and qualitative) needed are also not readily available and generally require a lot of additional research and input from many sources. Further, the types of skills needed for this kind of work are unusual, ideally spanning both traditional evaluative skills with innovative ones and a keen sense for stakeholder facilitation and a learning orientation, not to mention a solid grasp of systems thinking concepts and technical issues relating to climate change. Few individuals or even organizations have this combination of skills.

Finally, we have no time to wait for this kind of learning on transformative change. The climate crisis is materializing now, with grave consequences for both humans and many other species. Although climate change interventions have been

underway for several years – and some longer- most are in the early stages of implementation, and it will be a long time before they are completed. It will be even longer before their influence beyond their (project) lifespans will materialize. We therefore have a significant timing challenge, where we can evaluate where we are now, using the available emerging evidence and cognizant of the aforementioned challenges and limitations, or we can wait years or decades to do so. Most of us would agree that as a society we cannot afford to wait and should seek to understand what we can on this topic as soon as we can, even with the limitations.

Thankfully, we do have a growing body of knowledge, practice, and expertise in this space. The experiences on evaluating transformational change from the three major sources of climate and environmental funding point to the potential such financing can have in moving systems and markets towards more low-carbon, climate-resilient development pathways. We have identified common factors that define transformational projects and programs. These include targeting critical issues at the systems level from the outset to achieve depth, scale, and longevity of change.

The evidence also suggests that transformational change often takes time and a concerted effort, and that it can occur in unpredictable and non-linear ways. Yet, well-conceived and timed projects that respond to country demand and address critical barriers can lead to transformational change, or incremental change that can over time support transformation, even if their scope is rather limited.

Now, in 2019, we are seeing how many organizations and individuals are focused on this issue of transformational change, and how several other evolving efforts are working to understand how to materialize positive transformation – as well as how to evaluate it. We are identifying opportunities for further synergy amongst these efforts and are exploring ways to support strategic collaboration both formally and informally. As part of this, we want to continue to advance our collective knowledge and skills around innovative methodologies (some known, and some likely yet to fully materialize) that can advance systems change evaluation expertise and practice, and to share our experiences and knowledge so that we can be stronger as a field and as a global community.

The opportunities for systematic learning on transformational change naturally extend to even bigger issues, including sustainable development. The growing focus on transformation and its natural relevance to our world's most complex challenges is inspiring and will undoubtedly lead to global learning that can advance our collective work and support a more prosperous, equitable, and sustainable future.

References

Bird, Neil, Yue Cao and Adriana Quevedo. 2019. *Transformational change in the Climate Investment Funds. A synthesis of the evidence.* London: Overseas Development Institute. https://tinyurl.com/yxrblktp.

Ceballos, Gerardo, Paul R. Ehrlich, and Rodolfo Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences* 114 (30): E6089-E6096. https://www.pnas.org/content/pnas/114/30/E6089.full.pdf.

CIF. 2015. *Climate Investment Funds: Accomplishments, Transformational Impact, and Additionality in the Climate Finance Architecture.* Washington, DC: CIF JOINT CTF-SCF/TFC.15/3 (Joint Meeting of the CTF and SCF Trust Fund Committees)

CIF, and ICF. 2018. *Evaluation of the Climate Investment Funds' programmatic approach. Final Report and Management Response*. Washington, DC: CIF. https://tinyurl.com/yxajqfm9.

Ferguson, Briony C., Rebekah R. Brown, and Ana Deletic. 2013. Diagnosing transformative change in urban water systems. Theories and frameworks. *Global Environmental Change* 23 (1): 264–280. doi:10.1016/j.gloenvcha.2012.07.008.

Fiala, Nathan, Jyotsna Puri, and Peter Mwandri. *Becoming bigger, better, smarter. A summary of the evaluability of Green Climate Fund proposals.* Songdo, South Korea: Green Climate Fund (IEU Working Paper 2). https://tinyurl.com/y4l2fx3o.

GEF. 2015. GEF2020. S*trategy for the GEF.* Washington, DC: Global Environment Facility. https://tinyurl.com/y57ewxvc.

GEF IEO. 2017. *Evaluation of GEF support to transformational change.* Washington, DC: Global Environment Facility Independent Evaluation Office. https://tinyurl.com/yxmnu3e9.

Gell-Mann, Murray. 2010. Transformations of the twenty-first century. Transitions to greater sustainability. In: *Global sustainability. A nobel cause*, (eds) Hans Joachim Schellnhuber, Mario Molina, Nicholas Stern, et al., 1-8. Cambridge, MA: Cambridge University Press.

IPCC. 2018. *Global warming of 1.5°C.* Geneva: Intergovernmental Panel on Climate Change. https://www.ipcc.ch/sr15.

Kania, John, Mark Kramer, and Peter Senge. 2018. *The water of systems change.* [Boston:] FSG

Consulting (https://www.fsg.org/publications/water_of_systems_change)

Kuhn, Thomas S. 1962. *The structure of scientific revolutions*, 2nd ed. Chicago: University of Chicago Press (International encyclopedia of unified science).

Lewis, Simon L., and Mark A. Maslin. 2015. Defining the anthropocene. *Nature* 519: 171–180. doi:10.1038/nature14258.

Meadows, Donella H., Dennis L. Meadows, Jørgen Randers, et al. 1972. *The limits to growth.* New York: Universe Books.

Moore, Michele-Lee, Ola Tjornbo, and Elin Enfors. 2014. Studying the complexity of change. Toward an analytical framework for understanding deliberate social-ecological transformations. *Ecology and Society* 19 (4): 54. doi:10.5751/es-06966-190454.

Puri, Jyotsna. 2018. *Transformational Change. The Challenge of a Brave New World.* Songdo, South Korea: Green Climate Fund Independent Evaluation Unit (IEU) (Learning Paper 1).

Roy, Joyashree, Petra Tschakert, Henri Waisman, et al. 2018. Sustainable development, poverty eradication and reducing inequalities. In: *Global warming of 1.5°C.,* (eds) Valérie Masson-Delmotte, Panmao Zhai, Hans-Otto O. Pörtner, et al. https://www.ipcc.ch/site/assets/uploads/2018/11/sr15_chapter5.pdf.

McPherson, Sam, Matthew Savage, Jessica Kyle, et al. 2019. *Evaluation of transformational change in the Climate Investment Funds. Final evaluation report.* Hove, East Sussex: Itad. https://tinyurl.com/y2ac3mnt.

Schellnhuber, Hans Joachim, Dirk Messner, Claus Leggewie, et al. 2011. *World in Transition. A Social Contract for Sustainability.* Berlin: German Advisory Council on Global Change (WBGU). (Flagship report). https://tinyurl.com/y6qzh6bh.

World Bank Group IEG. 2016. S*upporting transformational change for poverty reduction and shared prosperity. Lessons from the World Bank Experience.* Washington, DC: World Bank Group (An IEG Category II Learning Product).

World Bank Group IEG. 2017. *China renewable energy scale-up program. Phase one. Project performance assessment report.* Washington, DC: Independent Evaluation Group of the World Bank Group.

World Commission on Environment and Development (WCED). 1987. *Our common future. Report of the World Commission on Environment and Development.* Oxford: Oxford University Press.

Systems evaluations for transformational change: challenges and opportunities

Cristina Magro and Rob D. van den Berg

ABSTRACT. In this chapter we argue that, if evaluators are to contribute to transformational changes required by the increasingly widespread global threats we are facing, they need to become fluent in systems thinking; to be open to evidence and sources of knowledge from various areas; identify, among the rich diversity of approaches, tools and methods available, the ones relevant and significant for their tasks; to provide insight and understanding on how interventions made through projects, programmes, and policies work contributing either positively or negatively to the dynamic equilibrium of systems. We provide an overview of some of the crucial foundations of systems approaches, exemplifying extensively from climate change and other areas. We formulate four challenges for evaluators and propose a way to contribute to further exchanges between science and evaluation with the support of IDEAS.

Introduction

While transformational change can happen at each level and scale of human activity, in nature or our environment, it is clear that the global crises that humanity is now facing call for major changes at higher and systems levels. There is, in fact, widespread agreement that a radical turnaround needs to happen on many issues. Let us take the climate crisis as an example. While many politicians have in the last three years been elected on platforms that included the denial of climate change (in the US, in Australia, in Brazil), there is nevertheless full agreement that our current energy systems and markets will collapse when fossil fuels run out, unless they transform in time toward renewable energy sources.

The sceptics of climate change object against funding climate action and against a quick move towards a transformation of the energy market. They claim that the costs are (too) high and the problems we are facing will be solved through technological innovation, and that no role for the government is required. Common statements are that current climate action is futile, for exactly the same reasons as many supporters of the Paris Agreement are increasingly worried about our planetary future: both sides perceive current climate action as insufficient to keep global warming within a 2 degrees Celsius limit. The sceptics and the supporters of climate action differ in their interpretation of what this means: the sceptics trivialise the dangers and express faith in technological solutions and the role of markets, whereas the supporters see evidence of catastrophic change and the need for an "all hands on deck" approach to transform societies, economies and our interaction with nature.

Evaluative evidence on climate action underscores the fundamental agreement on the promise and futility of current efforts to bring the world back from the brink of disaster. In 2013, in the First Report of the Fifth Overall Performance Study of the Global Environment Facility, its Independent Evaluation Office asked attention for the deficit for climate and environmental funding to make a dent in the global challenges of climate change, environmental degradation and biodiversity loss. According to calculations of the World Bank, funding levels should at least be a factor 10 higher to meet the needs. At the same time, funding going to unsustainable use of natural resources as fuel and water was at least 100 times higher (GEF 2013, p.2, para 22). This led one of us to the formulation of the micro-macro paradox in climate funding: while investments battling climate change tended to be successful, no such success was visible at the macro level of ongoing global heating (Van den Berg 2017).

The micro-macro paradox calls for broadening the boundaries of the systems considered, and emphasizes the need to design projects, programmes and policies in various realms with the full understanding of the broader context that may undo any good that these interventions manage to achieve. The micro-macro paradox points out that we should not be satisfied with the achievements of a single project, programme or narrowly defined policy. Evaluative evidence on the insufficiency of climate and biodiversity related development funding in the first two decades of this century has been synthesized by the Multilateral Development Banks and UN organisations for both the Convention on Biological Diversity (ECG, GEF, UNDP and FAO 2010) and the UN Framework Convention on Climate Change (ECG 2011). This could be likened to "preaching to the converted", and it should not be a surprise that these calls for evidence-informed action and funding went unheeded by governments throughout the world.

What certainly cannot be called "preaching to the converted" is the recent report of the Global Commission on Adaptation (2019), presented to the world, painting a dire picture of suffering and inequality if climate change continues without counteraction. It proposes investments of US$1.8 trillion to generate benefits of more than US$7.1 trillion in avoided losses due to climate change, economic benefits and social and environmental benefits. While this is revolutionary and transformative in its ambition and approach, it begs the question that if the world has not been willing to invest in preventing climate change, why would it be interested in invest in adapting to climate change? However, it adds an authoritative voice supporting transformational change, and establishes once again that the world is at a crossroads and evolution and incremental change will not sufficiently help us.

For evaluators to contribute to understanding how projects, programmes and policies can induce and support lasting systems changes, we need to become fluent in systems thinking, include systems analysis in our evaluations, learn and understand how systems operate, can be understood, studied and evaluated, and changed. In this chapter we will identify systems as composed of interrelated components in such a way that the functioning of the whole is bigger than the sum of its components. It is crucial that the systems which are object of projects, programmes and policies be integrated in the forefront of evaluations so that an entire evaluand system be considered, composed by the interventions triggered by the projects, programmes or policies with their peculiar interactions with the object of concern (as recently argued in IIED 2019). It is also crucial to recognize that systems are subject to positive and negative feedback loops that tend to balance each other out, and that the interventions need to contribute to break through unsustainable "balances" so that the system can transform itself in a sustainable direction.

Let us here also pay tribute to some of the many pioneers who have explored systems in evaluation in the past decade, knowing we cannot exhaust the list: Michael Quinn Patton (2010 and more recently with the Blue Marble Evaluation concept), Richard Hummelbrunner (2011), Bob Williams (2011, 2015), Aaron Zazueta and Jeneen Garcia (2014), Michael Bamberger, Jos Vaessen and Estelle Raimondo (2015) and Emely Gates (2017), amongst others, and build on their work.

This chapter aims at raising five challenges evaluators face in the light of the need for transformational change and will conclude with issues for further discussion in our profession.

BOX 1. Evaluators in the garden of systems

When discovering the world of systems thinking and analysis, evaluators will enter the gate into a brand-new world. They will encounter methods of analysis very different from what they are used to. Development evaluators are used to a toolbox overwhelmingly filled with sociological and economic instruments. On the quantitative side, their toolboxes have an emphasis on before/after, with/without comparisons based on field data that are analysed with frequency statistics and various regression techniques designed to prove statistical significance. On the qualitative side, the traditional toolkits contain sociologically inspired case studies and stakeholder interaction, including new analytical tools that can derive causal relationships from case studies, such as Qualitative Comparative Analysis (QCA).

They will suddenly enter an energetic and vibrant world that works with Bayesian statistics, where predictions take over from gathering large numbers of data, where modelling and scenarios are all the rage and achieve unexpected vistas, where observable reality has both chaotic and ordered characteristics, where tipping points can transform systems into something that did not exist before, and where phenomena seem to require the use of mathematical power laws, para-consistent logic, non-linear reiteration, fractals and Fibonacci sequences. On top of that, our evaluator is surprised to see this bright garden populated by a huge number of people they never met before: systems analysts, weather analysists, defence specialists, insurance financiers, investment bankers, traffic coordinators, logistical managers, and so on. In a distant corner of the garden they discover a group of macro-economists, who use systems techniques in their considerations of fiscal and monetary policy, but have been banned from general discussions of development policy to this garden by the micro-economists who hold sway in policy discussions about economics, evidence based policies and measurement of impact through randomized controlled trials.

How to begin to make sense of this lively new garden, and how to start growing flowers and plants that tackle the road towards achieving the Sustainable Development Goals is a big challenge evaluators face. It requires more than scholarly knowledge, like principles, skills and attitudes, and a willingness to consider and explore how systems behave according to their own structural possibilities and history.

A brief exploration of systems thinking

While systems thinking and analysis are often considered relatively new branches of knowledge, the roots of the word "system" (in Greek systema: an organized whole, a whole compounded of parts[1]) indicate that systems thinking has been part of human experience for a long time and is acknowledged in many cultures all over the world. One could argue that systems thinking has been intuitive throughout the world in religions and world views. While especially philosophical, religious and cultural perspectives may have run deep and sophisticated, the application of this understanding to societies, economies and our relationship with nature was often fraught with difficulties and challenges.

Systems thinking is not always experienced the same way for all. If for some it is natural and intuitive, for others it may sound clumsy and just not right. As Frakes and Linder note (2011, 2):

> […] even with repeated exposure, some individuals within any group consistently disregard the majority of systems thinking practices, dismissing their value or return on investment. In contrast, within any group, some individuals quickly embrace and delight in discovering the "language" of systems thinking. (In our classes, these are the students who say, "This is the way I've always thought; you're just giving me language to express it.").

Adopting systems thinking is often described as equivalent to learning a new language, with all richness this experience involves. Learning a language introduces us to different ways of living, thinking, reasoning and of creating worlds which may initially be misunderstood by us but turn out to be perfectly suitable for their inhabitants. To be successful in such an endeavour, it is necessary to put aside ingrained frames of mind and frameworks. To learn a language, in this sense, is genuinely different from learning a fixed set of words organized by a formulaic grammar which is supposed to convey fixed meanings. Similarly, to learn the basics of systems thinking is easy. To fully apply the concepts and the conceptual machinery, and to embody the attitudes of a systems thinker is not. It requires experience to become good at.

1 See https://www.etymonline.com/search?q=system.

Evaluators who are tasked to think about the systems that need to be transformed in order to achieve the Sustainable Development Goals, and are asked to evaluate an intervention implemented through a project, programme or policy in this light, may think: Do I feel comfortable in "systems thinking"? Is it easy to recognize, delimit and describe systems such as governance, institutions, markets, formal and informal sectors, ecosystem services, social and cultural systems, equity and gender structures and so on? Where can one learn more about systems thinking and build a bridge to evaluation? Could young and emerging evaluators gain an advantage by including these ideas in their academic studies and early practice?

Overall, these questions translate into the following challenge:

CHALLENGE 1 Evaluators need to become fluent in systems thinking to apply systems concepts, approaches and methods in their evaluations.

Evaluators need to become adept at systems thinking, as it is the start of applying systems analysis. In the past, methodological enrichments in evaluation approaches have taken place through including new expertise in evaluation teams, often hired ad hoc and for specific evaluations only. The shift that we advocate is more permanent and asks for practitioners to broaden their view – to consider larger units (or systems), in a nutshell. Working in multidisciplinary teams is one of the ways to do this. The challenge is to move beyond disciplinary pride and build conversations that are transformative in themselves.

The need to adopt systems thinking may provide an opportunity for the profession to embrace young and emerging evaluators, as many come fresh from universities with additional and new skills and theoretical and scientific perspectives. Their inclusion in evaluation practice may lead to the revolution of evaluation that is raised in chapter 3 by Bianca Montrosse Moorhead et al.

Systems analysis

In the history of science, systems thinking played an important role. One of the first modern great scientific revolutions was the paradigm shift at the end of the Middle Ages in Europe from an Earth-centred universe to a perspective in which gravity commanded the movements of the planets around the sun. This led to one of the first mathematical efforts to calculate the functioning of a system (i.e. the solar system). However, it did not lead to the development of science with a systems' view. Instead, a gap started to develop over time between approaches that would

isolate phenomena to study, and approaches that would aim for understanding and studying objects in their interconnectedness with other objects. And while some systems related sciences went from strength to strength (astronomy, geology, archaeology) and others included both micro- and macro perspectives (physics, evolution, social sciences), gradually the emerging prevailing perspective on scientific methodology became linear and reductionist in nature.

Especially in the philosophy of science, a tradition took the shape of reductionism[2] and a focus on observable phenomena. When the global empire of Great Britain and the economic power of the USA became dominant, reductionism (here "the scientific attempt to provide explanation in terms of ever smaller entities") became widespread and strong throughout the world. This was also encouraged by developments after the Second World War, where one kind of logic, reasoning and experimentation aiming at prediction, control and domination – of nature, of the world, of "undesirable" features and occurrences would achieve wide-ranging power and privilege. Strict and disciplined linear analytic methods, designed to isolate parts of concerned objects or phenomena and then dig into their essence, would lead to an understanding of the mechanics of the world as stable, immutable, ruled and precise as a Swiss watch.

Amazing accomplishments were produced in western science within this frame of mind. As an example, we may point out the Human Genome Project (HGP), international scientific research initiated in 1990 to determine the DNA sequence of the entire human genome. Success was announced in 2003, and in 2006 Nature published the last remaining uncharted chromosome, thus completing the project. Over the years, and through its reductionist approach, the project influenced areas as agriculture, animal husbandry, bioarchaeology, anthropology, and others. A potential big reward envisioned was the capacity to identify and correct future undesirable characteristics of a person or just any living being linked to particular features of the DNA – from cancer to alcoholism, criminal tendencies and mental diseases, reproductive capacities, etc. Ideas based on essentialism, straightforward causality and instruction of a system from the outside, as the ones that grounded the HGP, had severe limitations, which were already pointed out before the project was underway (see Oyama 2000 for a perspective beyond nature versus nurture).

2 For a good if somewhat journalistic overview of reductionism see https://en.wikipedia.org/wiki/Reductionism.

Philosophical and scientific reductionism focuses on underlying mechanisms that "cause" whatever happens to the systems as a whole, which led to the hope that if only we would know "what works" at the micro-level, we would be able to change the world. This found its expression in the evidence-based movement, focused on a reductionist vision of science, advocating a hierarchy of evidence, with the randomized controlled trial at the very pinnacle. The Maryland Scientific Methods Scale is exemplary for this approach, basically only admitting evidence that is based on counterfactual studies, ranging from observational to randomized controlled trials. Following the "success" of randomized controlled trials on medicine, the evidence-based movement aimed to introduce similar trials and studies in development and in social and economic policy formulation.

However, micro-level trials in development studies will not lead to the discovery of mechanisms that will change and transform societies, as "what works" at the micro level is not "what works" at the macro level. Let us give the example of a car: a well-functioning system of a steering wheel, a combustion engine, seats, doors, a gas tank and so on. Reducing the car to its components, we can see a lot of causalities that support the driving of the car: the connections of the steering wheel to the front wheels of the car, enabling it to change direction; the gas pedal ensuring that the combustion engine works harder, ensuring a higher speed of the car, and so on. However, no reductionist explanation of a car would be able to explain why it moved from A to B, as this is the decision of the driver and not linked to any specific component of the car. Driving the car from A to B is "caused" by a mechanism of the car as a whole and including a driver: preferences of the driver as far as destination is concerned, whether the roads are open or obstructed, if there is a petrol station halfway through and so on. And at the level of "traffic" we can discern many cars moving together, causing traffic jams, or storms and mist intervening, or an Earthquake which suddenly destroys the road. And the wider we scope on related systems, we see mobility issues, trade routes, infrastructure opening up areas for commerce or for living, and we may see areas opened up for commercial deforestation, thus contributing to climate change, or roads reducing habitat for species, leading to dramatic reductions in biodiversity.

While reductionism ruled many disciplines and many research efforts, systems analysis flourished in other disciplines, especially in the second half of the 20th Century. Earth sciences increased our understanding of our planet, including the discovery of the movement of tectonic plates and a deeper understanding of the geology of the Earth crust as well as our increasing knowledge of the past through archaeology and the history of evolution as witnessed in fossils. Other sciences deeply ingrained in systems thinking are astronomy, ecology,

technological sciences in general and transdisciplinary approaches like cybernetics and information theory.

A science based on systems analysis that the general public has heard about is climate science. The scientific consensus on what climate change is, how it is caused, and how it can be prevented, is based on modelling scenarios that look at levels of greenhouse gasses in the atmosphere. This modelling is transdisciplinary – it combines findings of sociological, economic, technological, chemical and meteorological research. It is strongly rooted in scientific understanding of climate processes – emphasis is put on the historical dynamic, on how parts of the system have interacted with each other in the past, through today, and into the future. This historical/longitudinal view replaces the static, photo-like approach, and patterns of behaviour are identified and taken into consideration. The Intergovernmental Panel on Climate Change (IPCC) is keeping track of scientific findings on climate change, including its political and economic risks and possible response options. The reports of IPCC have been accused by some of being too conservative in their estimates and by others of being too dramatic in their conclusions – and, so far, the general direction of the reports has been consistently confirmed in climate trends throughout the world.

The information revolution that the world is experiencing since the Second World War is increasingly approached through systems analysis, with Big Data, Artificial Intelligence, and changing social, economic and technological perspectives due to internet and mobile connectiveness. The introduction of new technologies is ever faster, and their consequences are often (partly) unforeseen. Informational technology and Big Data analysis tend to use mathematical tools that are relatively under-utilized in development and in developmental evaluation, such as Bayesian statistics, risk analysis, Artificial Intelligence algorithms and so on.

In this regard, it is important to note that mainstream development and mainstream evaluation use primarily traditional tools, not designed to deal with complex and dynamic phenomena. This is especially hampering randomized controlled trials, as they need large numbers of data in order to be able to conclude on statistical correlation and significance, thus making randomized controlled trials more expensive and difficult to implement and less relevant to existing practices, as only static processes can be studied.

Bayesian statistics has long been anathema in science as it uses existing knowledge as input in statistical analysis, which then is refined through iterative processes to indicate conditional probabilities. While in science Bayesian statistics after the

Second World War was ostracized, it refused to die out in the world of applications: defence, heuristics, insurance, risk analysis, operations research and so on. An excellent overview of how Bayesian statistics overcame its condemnation by mainstream statistics can be found in Sharon Bertsch McGrayne's *The Theory That Would Not Die* (McGrayne 2011).

While many evaluators of development continue to focus on frequency statistics in their analysis of data, the mathematical applications for understanding our world and our reality have grown dramatically. In development, in general, the identification of risks is reduced to adding some assumptions of what could potentially go wrong in the logic of intervention in either a theory of change or a log frame. Risk analysis, in its more structural incarnation, is a quantitative approach to risks associated with scenarios of development, and thus not linear but multi-dimensional. The probability of different potential outcomes is calculated and can then be used for management purposes. This way of reasoning and approach has become mainstream in the insurance industry and in financial investments, with for example hedge funds calculating how much they should invest in different scenarios in order to hedge their funds against loss in one scenario versus gains in another scenario.

The bestseller author Nassim Nicholas Taleb, a former trader and risk analyst, turned publicist and philosopher, has frequently asked attention for mistaken use of statistics, especially warning for potentially catastrophic "black swan" events: focusing on unforeseeable events that demonstrate that a theory or assumption is not correct, often with dramatic consequences. He argues that these black swan events may be hidden in the regions of lower risks – the tail ends of calculated risks that exhibit a Gaussian distribution (Taleb 2018). He makes a distinction between risks that are shared (where low risks can be taken as acceptable for the average person) and risks that are personal (where even a low chance for disaster is certain to be fatal in the longer run). One could identify climate change as a risk that is "personal" for humanity and the planet. Mathematical risk analysis is now brought to bear upon investment decisions. The Network for Greening the Financial System (2019) has recently published a report on climate change as a source of financial risk.

Risk analysis from a systems perspective is not yet used in evaluations, as it may often be perceived as dealing with the future instead of the past, and evaluators are reluctant to address the future. But if science and transdisciplinary approaches increasingly focus on the future as worthy of our concern, evaluators should step over this threshold and start to include whether interventions that are supposed to reduce risk over time, in fact do so. And they should have the means at their disposal to judge this.

Adaptation to climate change, or climate change resilience may be a good example. Policies and interventions that promote these are supposed to lead to lower risks in the future, or lower costs for risks that cannot be prevented. This should by necessity lead to shifts and changes in the risk scenarios calculated for the future, and evaluators should include this in evaluations as their judgments on past performance require inclusion and understanding of risk analysis regarding the future. The Global Commission on Adaptation (2019) has done this on a global scale; are evaluators enabled to do this on a national or local scale?

Development interventions tend to look at the world, our societies and economies as inherently stable, and the intervention as a gentle or sometimes forceful push in the right direction, with outcomes often in a linear progression from the inputs. Sustainability is still perceived by many as a continuation of benefits after an intervention has ended; in other words: a new stable situation has been reached. However, stability in systems is dynamic, and changes in systems are non-linear.

Stability and change in systems that operate around us tend to be difficult to predict, as too many causal influences happen at the same time, in different directions. The effort to identify how natural systems interact with human systems through the Millennium Ecosystem Assessment (Millennium 2005) demonstrates how complex the various linkages between systems are, whereas more recent research demonstrates that all Sustainable Development Goals are interlinked, and that this needs to be taken into account to understand how to achieve goals without undermining others (International Council for Science 2018 and Scharlemann 2016).

While everything seems to interact with everything, the world "seems" relatively stable to us, and transformational change seems very difficult to achieve. Systems analysis recognizes that systems may be in equilibrium and tend to describe this as a whirlpool of vectors of change that mostly cancel each other out. And when a tipping point is reached at which a system changes into something different, this may be unexpected.

The class of mathematics that is used to study non-linear, contingent and historical phenomena that could be described as systems transforming is called "power- law mathematics": one quantity varies as the power of another ($y=x^k$). Variations of this include chaos mathematics, fractals, iterative number series such as the Fibonacci sequence (where each new number is the sum of the two previous numbers), and so on. The systems that we want to transform to achieve the Sustainable Development Goals, when measured on key data, tend to show power-law phenomena. This looks promising, as we will be looking for the wings of the

butterfly (and a small investment may lead to great changes), but it also poses great risks and uncertainties, as the hurricane that is caused by the wings of a butterfly may go in quite a different direction and destroy what we cherish.

An example of the need to study power laws in development is offered by Andriani and McKelvey (2007) who argue that management research needs to shift from Gaussian statistics toward extreme events and power laws, as this will reflect the real world better than the stable numbers of Gaussian averages. While they recognize the value of Gaussian statistics in situations where stability seems to reign, they also claim that this closes our eyes to non-linear change and transformation. Statistics in evaluation is often low-key if not hidden in annexes. Power-laws are rarely used in evaluations. If evaluations are to study transformation and help and support us to transform in the direction of achieving the Sustainable Development Goals, our profession will need to incorporate an understanding of how power laws predict changes.

As evaluators, we face the challenge of including the multitude of systems analytics and its achievements in our evaluation toolbox. While evaluations regarding climate action and other interventions that include the relationship between humanity and its habitat have exhibited various efforts to integrate systems perspectives in their toolbox, other areas of development evaluation have not yet proceeded accordingly. Transformative evaluation, whether at the gender and community level or at levels of society (as exhibited in chapter 2 by Osvaldo Feinstein) poses a major issue to the evaluation community, which leads us to our second challenge.

CHALLENGE 2 Evaluators need to be open to evidence and sources of knowledge generated through systems analytics, including especially future scenarios and power-law phenomena that lead to transformational change.

Introducing systems thinking and analysis in evaluations

While there is general agreement on the need to shift to systems to understand and help promote transformational change that will support the achievement of the Sustainable Development Goals, there is a lack of agreement or consensus on how this should be done. Just like systems thinking, evaluation is a field informed by various disciplines since its inception. The renewed interest and growing resource to systems thinking in evaluation faces immediately the challenge of understanding the implications of using systems approaches in our field:

> The theories, concepts, methodologies, and tools that are central to these fields often have multiple, and sometimes incompatible, definitions and numerous variations (Henning and Chen 2012). As evaluators select and translate systems approaches, they will be defined and applied differently to be relevant to and useful for evaluators. Finally, evaluators drawing on the systems fields are doing so within different evaluation theories (e.g., theory-based, responsive, equity-focused) for evaluating different kinds of interventions (e.g., social innovations, health care policy, regional development), and in different circumstances. Consequently, it is challenging to generally examine implications of systems approaches for evaluation practice. (Gates 2017, p.153).

Three reasons motivated Emily Gates to develop a study on the implications of systems approaches for evaluation practice:

1. New kinds of social interventions being developed by evaluation commissioners and stakeholders (e.g., networks, emerging innovations, and systems change), which require adequate evaluations;
2. The now relatively common claim, by evaluators, that they are making use of systems approaches (Patton 2016), without empirical research examining the use of these ideas and methods in practice;
3. A growing interest in research on evaluation.

Although we largely agree with these questions, our concern goes beyond emphasizing the instrumental role of evaluation, underscoring learning in the world of projects and programs itself – an emerging characteristic of evaluators' practice identified by Gates (2017). Besides designing evaluations which are adequate to the interventions promoted by projects, programs, and policies and to help improve them, it seems urgent that evaluations contribute to develop a good understanding of the systemic dynamics of the phenomena/systems which are object of these interventions, which often pose problems that cannot be solved with the very frame of mind that produced them – to refer to Einstein's famous quote: "The significant problems we face today cannot be solved with the same level of thinking we were at when we created them."

The question is then how to integrate systems perspectives. Williams (2011) recalls an inventory of systems approaches built in 2002 by the International Institute for General Systems Studies (IIGSS), which came up with more than 1200 approaches, some obscure, and others very well known. "No wonder," writes Williams (2011, p.2), "some people find it difficult to sort through the breadth and diversity of approaches to determine what is useful to them and what is not."

CHALLENGE 3 Among the rich diversity of approaches, tools and methods that systems analytics offer, evaluators need to be able to identify the ones relevant and significant for the task they are about to undertake.

Our suggestion is first to adopt a common definition of what systems are, second identify which systems play a role or are the focus of the intervention you are supposed to evaluate and third bring in specific systems science and analytical tools appropriate for the systems identified, as well as understanding of the negative and positive feedback loops affecting these systems.

Identification of systems

We start by defining *system*, describing its characteristics and formally summarizing the *types of systems* already mentioned elsewhere in this chapter. Needless to repeat what has been said above about the variety of ways in which these concepts are defined – the definitions below are our choice.

> *Systems* are dynamic units that we distinguish and choose to treat as comprised of interrelated components, in such a way that the functioning of the system, that is, the result of the interactions between the components, is bigger than the sum of its components.

One could be surprised by the reference to the "choice" of the observer to identify something like a system or treat it as an object in itself. An example may clarify this: you can choose to treat your watch as a simple unit, that is, not refer to how many and which components it has inside and use it as a device that gives you the time. Or, either out of curiosity or due to damage it may have suffered, you may choose to or need to open it or to take it to someone who can check how its components are and how they are interacting with each other.

A second issue to be highlighted is that systems can and will be composed of other systems. To go back to the example of a car previously explored: while the car is an easily recognizable system, the combustion engine in the car (or increasingly the electric battery!) can also be recognized as a system in itself, within the system of the car. Again, the behaviour of the system depends on what you choose to consider: while a combustion engine is a great system for getting movement out of burning fuel, and this translates upwards to the car, the ability of the car to move from A to B does not translate down to the combustion engine in itself. It lacks wheels and a steering mechanism.

A third aspect to bear in mind is the dynamics of systems. In fact, a great deal of the systems theories aimed to tackle this fundamental phenomenon that needs to be observed and dealt with, in its complexity and contingent history, especially in natural and mixed systems. Systems interact, all the time, with other systems that can even be distant in time and place. While identifying systems, it is paramount to recognize that they are not operating in a void – there will always be dynamic interactions with other, related systems.

For exploratory purposes of this chapter, we consider the following types of systems:

1. Designed systems – All machines are designed systems. Databases, legal systems, inventories of all kinds are equally man-made.

2. Natural systems – climate, forests, galaxies, human bodies, cells are all examples of natural systems, with their special characteristics regarding causality and dynamics explored above through the climate examples.

3. Mixed systems – society is the canonical example of mixed systems, exhibiting characteristics of natural and of designed systems.

Depending on the kind of questions we need to solve in our evaluations, other typologies can be useful. When what is at stake are the kind of components involved, it is possible to distinguish between:

1. Hard systems – made of concrete components.

2. Soft systems – like legal, institutional, religious, cultural and art systems, that may involve concrete components but overall sets of values, beliefs, principles, rules, etc.

3. Mixed systems – composed of soft and hard units/systems.

Systems related to the evaluand

A first exploration of systems may lead to a feeling that "to deal with wholes", "to enlarge the view" and to embrace objects in their circumstances is impossible to accomplish. Some practitioners of different areas of work have reported anxiety and overwhelming pressure that seems insurmountable when looking for the whole in a complex and interrelated world. Systems thinking has been frequently called *holistic thinking*, commonly understood as a mode of reasoning to deal with complex "wholes". While the use of this expression highlights the appeal to observers (scientists, evaluators, ourselves, etc.) to develop awareness by seeking to understand a larger picture than the one they are used to, on the other hand it

may lead to the search for "wholes", no matter how large they are, in a helicopter view. It soon becomes evident that this is not operational, and that flexibility needs to be maintained in order to understand and study systems.

Another difficulty is to accept that the complexity involved in systems nested within systems, and interactions between systems, such as between the social, economic and environmental domains, is not mechanistic as in the reductionist perspective on reality. The economist W. Brian Arthur (1999), an authority on complexity in economics and technology, described it as follows:

> Complexity is looking at interacting elements and asking how they form patterns and how the patterns unfold. It's important to point out that the patterns may never be finished. They're open-ended. In standard science this hit some things that most scientists have a negative reaction to. Science doesn't like perpetual novelty.

This "perpetual novelty" is what makes systems thinking and analysis a voyage of discovery and gives us hope that evaluators may contribute to providing evidence on what is happening in systems that could be used as a starting point for new directions.

Systems and especially nested systems can go up all the way into space and a good example is the global effort to repair the ozone layer that protects us from harmful rays of the sun. The thin layer of ozone in the lower portion of the stratosphere of Earth, that performs this essential service for our planet, was harmed and broken by substances in the air that were introduced to solve problems in refrigerators. This led to an emerging ozone hole that shook the world into action. The system of breaking down the Ozone Layer was stopped through a series of interventions that stopped production, forbade consumption (use in refrigerators) and organized the storage and ultimate destruction of Ozone Depleting Substances taken out of circulation. The choice of system should thus lead to the question of how this can be dealt with in an evaluation.

Complex systems may have mixed characteristics. This has led to the well-known starting points for identification of the system you want to evaluate:

1. To establish the boundaries of the system;
2. To identify the actors/components of the system;
3. To identify their relationships.

We add a fourth one that is essential to understand what is happening:

4. To identify positive and negative feedback loops that change the system, keep it in balance or move it in an unsustainable direction.

In evaluation we might always deal with mixed systems: besides the intervention examined, canonically called "the evaluand", which is a series of interventions in a project, programme or policy, the evaluator himself needs to be considered part of the system, together with stakeholders in general, and all values, frames of mind and beliefs that are permanently present in the activity. Chapter 6, by Hur Hassnain and Inga-Lill Aronsson, in this book, calls attention to the importance of heritage to the recovery of peace in fragile and conflict areas, bringing light to one component of social dynamics that, although immaterial (not concrete, soft – to use the term selected previously) plays a relevant and recognizable role in social identity and cohesion.

The dynamic of systems is driven by opposite forces – feedback and iterative loops that go in different directions. One direction goes to change: for the better or for worse. The other is feedback to restore the balance. For example, in climate change we are experiencing feedback loops that strengthen and reinforce climate change, such as melting of ice caps and glaciers that speeds up warming of the planet. At the same time, other feedback leads to natural sequestration of CO2 that will bring the CO2 in the atmosphere back to the levels Earth experienced before humanity started to use fossil fuels on a global scale. Natural processes take CO2 out of the air and deposit it in soil and at the bottom of the ocean, eventually turning it into rock. For our immediate future this is quite problematic, as these processes will take about half a million years to get us back to pre-industrial levels.[3] And as we have seen, the negative feedback loops for climate change far outweigh any positive ones at the moment, even if initiatives have started to speed up the natural positive feedback loop, for example through reforestation. Similar examples can be given in relation to other SDG goals, for example on the complex relationship between education, health care and livelihoods, where small changes in one area may lead to emerging new feedback loops that work beneficially or actually make goals harder to achieve.

Given these positive and negative feedback loops, any equilibrium of systems is therefore a dynamic equilibrium. It is historical, contingent with the interactions of the system with its surroundings and with the changes in related systems. The fact that a system appears stable, or appears to head in a certain direction, does not mean that we know how to control the path of its history or to fully predict its behaviour.

3 See https://www.nature.com/articles/climate.2008.122.

Choice of appropriate analytical tools

We need a systems perspective because we need to give birth to and support changes on unprecedent scales and in different realms; because we have a history of interventions of all kinds in various areas that have not returned the effort and the money invested the way we expected; because we have plenty of examples in our own life span that gives us evidence of interrelations between phenomena of various natures we were thought not to acknowledge by a long tradition of linear and reductionist thinking; because our traditional science and epistemology are good to deal with static situations, and we aim for a better and dynamic future learning from our history. Our systems need to change from the current situation they are in – they need to transform for humanity to have a sustainable future on our planet, in equitable and inclusive societies, leaving no-one behind and in balance with the resources of our planet.

One way to look at this has been the effort to establish humanity's global footprint, calculating how humanity's demand for ecological resources and services relates to the resources that planet Earth can generate in a year. The Global Footprint Network tracks resource use each year and identified 29 July 2019 as "Earth Overshoot Day", when more resources were used since 1 January 2019 than Earth can renew in a full year.[4] It is clear this cannot continue. This view is at the root of Greta Thunberg's activism and her "school strike for the climate". Greta has claimed that she is only repeating on the urgent need for climate action what scientists "have been communicating to the public for decades". This is correct. Now that the need for climate action is becoming more urgent than ever, the call for transformational change is becoming widespread. It is not unexpected that the climate and environment funds (see Uitto et al., chapter 7) have been most active in exploring evaluation of transformational change at global, national and societal levels.

Climate change used to be an energy issue first and foremost, and a solution would have been to shift our energy sources from fossil fuels to renewable energy sources, such as solar, wind, geothermal, and so on. The current climate crisis is a cross-cutting issue throughout the Sustainable Development Goals. Any call for transformational change for the Sustainable Development Goals therefore needs to include climate perspectives and the overuse of our ecosystem's resources. And vice-versa, any call to stop climate crisis needs to include contributions from other realms

4 See https://www.overshootday.org/.

in which the Sustainable Development Goals have their goals established. Many interrelations among phenomena of the various Sustainable Development Goals, apparently disconnected and distant, are available to be explored and used towards lasting transformational changes towards a sustainable and inclusive future.

The systems level that evaluators need to aspire to is therefore high and comprehensive. It is worth noting that size, in itself, does not guarantee a systems view which requires particular ways of dealing with dynamics, interrelations and history. In science, geologists, geographers and historians have studied encompassing systems of interactions between humanity and the planet. It has led to the identification of the age of mankind, the Anthropocene, as a new geological era on planet Earth.[5] Historians have looked at how the Anthropocene has taken shape over time in the influential book *The Shock of the Anthropocene* by Christophe Bonneuil and Jean-Baptiste Fressoz (2013) and the geological perspectives have been thoroughly discussed in the seminal publication *The Human Planet* of Simon L. Lewis and Mark A. Maslin (2018). The latter book provides a complex system perspective on how humanity has interacted with our planet, focusing on the role and availability of energy, with each era moving towards higher levels of energy use, going from hunter-gatherers, to the agricultural revolution, to mercantile capitalism, to the industrial revolution, to consumer capitalism. Their integrated, systemic and high-level perspective also includes a perspective on transformational change. They warn about the inherent resilience in systems and that the "breach of multiple planetary boundaries" (Lewis 2018, p. 349) that we currently experience is not easily changed, as our current "system" based on fossil fuels, hybrid crops and nitrogen fertilizer is dynamically resistant against the transformational change we need and hope for.

Recognition of positive and negative feedback loops

In line with our earlier discussion of the dynamic nature of systems, Lewis and Maslin (2018) describe the highest-level order of our societies and economies, in their relationship with nature, not as stable and resistant to change, but as instable and moving in an unsustainable direction. Our challenge is not just to support and evaluate action towards a sustainable future, but to do so in light

5 The first step towards formal recognition of the Anthropocene as a new geologic epoch has been taken in May 2019 by a panel of scientists, preparing for a decision in 2021 by the International Commission on Stratigraphy.

of transformations that are happening that undermine our current societal and economic arrangements. It seems that while our economies have gained income and wealth, both on the social and the natural side transformations have been in unsustainable directions, perhaps more so on the natural side, with three major environmental crises that threaten humanity's habitat. The first is climate change, which is well known and forecast to be on a path to irrevocably change our climate. This transformative change in the wrong direction may over hundreds of years lead to regions in the tropics becoming inhospitable due to high temperatures, may melt all ice on the planet, leading to sea level rise with tens of meters, leading to large scale resettlement of coastal populations and cities to higher inland and leading to dramatic changes in agriculture and food availability throughout the world.

The climate crisis is not the only one that displays weak sustainable and strong unsustainable feedback loops. A second environmental crisis concerns the loss of biodiversity, now often recognized as the sixth mass extinction of life on the planet, threatening humanity's survival (Kolbert 2014). This loss of biodiversity will disrupt our agricultural systems and will lead to food crises that are unprecedented. Globalisation is at the root of extinction, as formerly isolated areas are increasingly opened up to invasive species and viruses and bacteria that potentially can turn ecosystems into wastelands. What happened to the American Indians when flu strains reached them that they had no resistance against, is now happening to ecosystems and species of plants and animals that have no resistance against invasive species and diseases.

The third crisis concerns the chemicals and waste that our societies let loose on the world. While plastic waste in the ocean may be somehow fished up, many chemical components endanger our ecosystems without being visible. One of the most remote islands in the world, Henderson island in the Pacific, has been visited in 2018 by a team of scientists who established that no less than 18 tons of plastic waste has been washed ashore on its formerly pristine beaches.[6] Plastic will break down to micro particles that have by now entered into the food chain and end up in our bodies. New chemicals and organic substances are entering industrial processes in our consumer societies. Despite assurances of the industries producing these for packaging and conservation purposes, traces of them will end up in our bodies.

These transformational processes need to be reversed, which is a tall order, as these processes all still receive large inputs to firmly continue them on their way to

6 See the Guardian of 30 July 2019 at https://tinyurl.com/y424vs5f.

transform our world. A major challenge evaluators will be facing is to value and judge efforts towards transformational change in the direction of a sustainable future in light of dominant change processes in the other direction. Are transformational processes supported sufficiently and adequately to reverse trends? Understanding is thus key. Evaluators will need to understand in which directions systems are moving, and whether a policy, programme or intervention is ensuring a transformational change in the right direction, or whether it is a lost race.

This leads us to a fifth and last challenge we will formulate here, for evaluators and for evaluation.

> **CHALLENGE 4** Evaluations need to provide insight and understanding of whether interventions and policies are able to reverse trends and may overcome barriers to move in sustainable directions.

Further discussion in our profession

When we as evaluators identify the systems we need to consider, understand and value in our evaluations, we need to start the search for systems science and tools that may help us. The quest may start by identifying ongoing related research, the scientists who undertake this and interact with them to see which insights we may use and rely on, which data sets are available for us and which analytical tools are handy. We will aspire to "dynamic evaluations" and, to promote learning and changes beyond immediate structures of the systems involved, in continuous and multiple loops of reflections, as described by Osvaldo Feinstein in chapter 2 in this book.

IDEAS is proposing, for further discussion in the global evaluation community, to host an International Evaluation Academy that may play an important role in facilitating the linkages and exchanges between science, research and evaluation on systems issues and regarding the Sustainable Development Goals. Moreover, the International Evaluation Academy may function as a catalyst for ideas, professionals, and resources to feed a global intelligence for the advancement of evaluation to fulfil its mission and the task of contributing to the achievement of the SDGs. This proposal is a consequence of IDEAS involvement with professionalization, including former initiatives as the Competencies Framework and the Code of Ethics, both elaborated by a committee of members from all around the globe, as well as our efforts to create thematic groups that discuss topics of interest to the global evaluation community.

This is not a practical guide to how systems analysis can be integrated into our evaluations. Our purpose is to start a discussion on how evaluation can include systems thinking and analysis to be significant and useful for efforts to transform our systems in the direction of aspirations of the Sustainable Development Goals, leaving no-one behind, building inclusive societies and safeguarding our planet against growing inequity, political populism, conflict and violence, as described from various perspectives in this book.

We hope that the four challenges we have formulated generate discussion in our profession at the Global Assembly and at other forums:

CHALLENGE 1 Evaluators need to become fluent in systems thinking to apply systems concepts, approaches and methods in their evaluations.

CHALLENGE 2 Evaluators need to be open to evidence and sources of knowledge generated through systems analytics, including especially future scenarios and power law phenomena that lead to transformational change.

CHALLENGE 3 Among the rich diversity of approaches, tools and methods that systems analytics offer, evaluators need to be able to identify the ones relevant and significant for the task they are about to undertake.

CHALLENGE 4 Evaluations need to provide insight and understanding of whether interventions and policies are able to reverse trends and may overcome barriers to move in sustainable directions.

References

Andriani, Pierpaolo, and Bill McKelvey. 2007. Beyond Gaussian averages. Redirecting international business and management research toward extreme events and power laws. *Journal of International Business Studies* 38 (7): 1212-1230.

Arthur, W. Brian. 1999. *Interview*, 16 April 1999. https://tinyurl.com/y2qdch6p.

Bamberger, Michael, Jos Vaessen, and Estelle Raimondo (eds). 2015. *Dealing with complexity in development evaluation. A practical approach.* Thousand Oaks, CA: SAGE Publications.

Bonneuil, Christophe, and Jean-Baptiste Fressoz. 2017. *The shock of the anthropocene. The Earth, history and us.* London: Verso. Transl. from the French by David Fernbach.

Evaluation Cooperation Group (ECG). 2011. *Briefing note. Overcoming barriers to energy efficiency. New evidence from independent evaluation.* https://tinyurl.com/yx9ees5j.

Evaluation Cooperation Group (ECG), Global Environment Facility (GEF), United Nations Development Programme (UNDP), and Food and Agriculture Organization of the United Nations (FAO). 2010. *Briefing note on ensuring biodiversity in a sustainable future. Lessons from evaluations.* [Nairobi:] UNEP (UNEP/CBD/COP/10/INF/39 14 October 2010). https://tinyurl.com/y4yj8m2y.

Frakes, Jeffrey, and Nalani Linder. 2011. A new path to understanding systems thinking. *Systems thinker* 22 (8): 2-7. https://tinyurl.com/yxld2b8e.

Gates, Emily F. 2017. Learning from seasoned evaluators. Implications of systems approaches for evaluation practice. *Evaluation* 23 (2): 152–171. https://doi.org/10.1177/1356389017697613.

GEF IEO. 2013. OPS5. *Fifth Overall Performance Study of the GEF. First report. Cumulative evidence on the challenging pathways to impact.* Washington D.C.: GEF Independent Evaluation Office.

Global Commission on Adaptation. 2019. *Adapt now. A global call for leadership on climate resilience.* Rotterdam and Washington, DC: Global Center on Adaptation and World Resources Institute

Hummelbrunner, Richard. 2011. Systems thinking and evaluation. *Evaluation* 17 (4): 395–403. http://evi.sagepub.com/content/17/4/395.

Hummelbrunner, Richard, and Bob Williams. 2011. *Systems concepts in action. A practitioner's toolkit.* Stanford, CA: Stanford University Press.

International Institute for Environment and Development (IIED) et al. 2019. *Briefing. From monitoring goals to systems-informed evaluation. Insights from SDG14.* (Briefing Paper 11, May 2019). https://tinyurl.com/yy3gmmbe.

International Council for Science. 2018. *A guide to SDG interactions. From science to implementation.* https://tinyurl.com/ybsxbazt.

Kolbert, Elizabeth. 2015. *The sixth extinction. An unnatural history.* London: Bloomsbury.

Lewis, Simon L., and Mark A. Maslin. 2018. *The human planet. How we created the anthropocene.* [London:] Pelican Books.

McGrayne, Sharon Bertsch. 2011. *The theory that would not die. How Bayes' rule cracked the enigma code, hunted down Russian submarines and emerged triumphant from two centuries of controversy.* New Haven: Yale University Press.

Millennium Ecosystem Assessment. 2005. E*cosystems and human well-being. Synthesis.* Washington, D.C.: Island Press.

Network for Greening the Financial System. 2019. *A call for action. Climate change as a source of financial risk.* https://tinyurl.com/y2m3f582.

Oyama, Susan, and Richard Lewontin. 2000. *The ontogeny of information. Developmental systems and evolution* (2nd edition). Durham, N.C.: Duke University Press (Series on Science and Cultural Theory).

Patton, Michael Quinn. 2010. *Developmental evaluation. Applying complexity concepts to enhance innovation and use.* New York: Guilford Press.

Patton, Michael Quinn. 2016. What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models. *American Journal of Evaluation* 37 (2): 250–65.

Scharlemann, Jörn P.W, Rebecca C. Mant, Nicholas Balfour et al. 2016. *Global goals mapping. The environment-human landscape.* Brighton (UK) and Cambridge (UK): Sussex Sustainability Research Programme, University of Sussex, Brighton, UK and UN Environment World Conservation Monitoring Centre. https://tinyurl.com/y269ns3u.

Taleb, Nassim Nicholas. 2019. *Skin in the game. Hidden asymmetries in daily life.* [London:] Penguin Books.

Van den Berg, Rob D., and Lee Cando-Noordhuizen. 2017. Action on climate change. What does it mean and where does it lead to? In: *Evaluating climate change action for sustainable development,* (eds) Juha I. Uitto, Jyotsna Puri, and Rob D. van den Berg, 13-34. Cham: Springer Open.

Williams, Bob. 2011. All methods are wrong. Some methods are useful. *Systems thinker* 22 (4): 2-7. https://tinyurl.com/yxubkw7l.

Williams, Bob. 2015. Prosaic or profound? The adoption of systems ideas by impact evaluation. *IDS Bulletin* 46 (1): 7-16.

Zazueta, Aaron, and Jeneen Garcia. 2014. Multiple actors and confounding factors. Evaluating impact in complex social-ecological systems. In: *Evaluating Environment in International Development* (ed) Juha I. Uitto, 194-207. London: Routledge.

Contributors

Josette Arévalo is a Board member of the EvalYouth Global Network and Founder of EvalYouth LAC and EvalYouth Ecuador chapters. In recognition of her contributions to the evaluation field, she received the Young and Emerging Evaluator award at the 2017 Global Evaluation Forum. Josette has a Ph.D. in Public Policy and Administration and an MPP at George Washington University, and a master's in development economics at Latin American Faculty for the Social Sciences (FLACSO), Ecuador. She has worked as an advisor to the Minister of Education of Ecuador, as the Executive Director of the National Institute of Educational Assessment of Ecuador, and currently is working as a university professor. josette2003@gmail.com.

Inga-Lill Aronsson is a cultural anthropologist and Senior Lecturer in museum and heritage studies at Uppsala University, Sweden. She is a member of the IDEAS Board. She has extensive fieldwork experiences foremost in Latin America. Her research interests are longitudinal studies on resettlement and heritage, memories, and reconciliation. She is the former Director of the Network on Humanitarian Action (NOHA). Her teaching, supervision and curriculum experiences are extensive and multi-disciplinary including a high-profile international master course on heritage and conflict. Her latest publication is on how to use the past in resettlement evaluation. She is at present engaged with an assessment of how to address bad qualitative participatory research. inga-lill.aronsson@abm.uu.se.

Lennise Baptiste has more than 30 years of combined professional experience in the fields of education and evaluation. She began the evaluation of projects and programmes in the United States in 2005 while employed at the Research and Evaluation Bureau at Kent State University, Ohio. In 2011 she returned to the Caribbean and has been working across the English-speaking countries evaluating projects for different donor agencies as an Independent Consultant. She has been a member of the American Evaluation Association since 2006 and was the Chair of the Mixed Methods Topical Interest Group from 2008-2018. She is the Immediate

Past Chair (2015-2018) of the Caribbean Evaluators International (CEI) Board and continues to work on behalf of the CEI on the South to South Evaluation Initiative which links the CEI to other Voluntary Organisations of Professional Evaluators in the Global South. lbaptiste.mepro@gmail.com.

Geeta Batra is the Chief Evaluation Officer and Deputy Director of the Global Environment Facility Independent Evaluation Office since June 2015. From 2010-2015 Geeta was the Manager of Country and Corporate Evaluations in the Independent Evaluation Group of the World Bank. Before that, she set up and managed the Results Measurement System for Advisory Services in the International Financial Corporation, the private sector arm of the World Bank, from 2004-2010. She was in the private sector development department of the World Bank from 1998-2004, responsible for investment lending and advisory operations, and at American Express, New York from 1996-98. She has conducted and managed project, program and country-level evaluations in a leadership role over the past 16 years. Geeta has a Ph.D. in Economics and an M.BA. gbatra@worldbank.org.

Arlene Bailey is a Senior Research Fellow at the Sir Arthur Lewis Institute of Social and Economic Studies (SALISES), University of the West Indies (UWI), Mona. Arlene's research areas include assessing the role of information and communication technologies (ICTs) and community informatics for sustainable development. Arlene currently serves as a Senior Editor for the Electronic Journal of Information Systems in Developing Countries and as an Associate Editor for Information Technology for Development. She also serves as Secretary of The International Federation for Information Processing (IFIP) Working Group 9.4 on the Social Implications of Computers in Developing Countries. arlene.bailey@uwimona.edu.jm.

Rob D. van den Berg is the President of the International Development Evaluation Association (IDEAS), in his second and final term. He is a Visiting Professor at the International Development Institute, King's College, London. He is also an Honorary Associate at the Centre for Development Impact at the Institute of Social Studies in Brighton, U.K. He is an adviser to the Climate Investment Funds, as well as a member of the Advisory Council of Wilton Park. From 2004 to 2014, he was Director of the Global Environment Facility Independent Evaluation Office in Washington, D.C. Before that, he was Director of Evaluation of the Dutch Ministry of Foreign Affairs from 1999 to 2004. He has edited several books and published more than 25 articles and chapters on development, research and evaluation. Rob has a master's in contemporary history at the State University of Groningen in the Netherlands. ideaspresident@gmail.com.

Khalil Bitar is the Vice Chair of EvalYouth Global Network; Founder of the Palestinian Evaluation Association (PEA); Co-Founder of EvalYouth MENA Regional Chapter; and Board Member of the MENA Evaluation Network (EvalMENA). He works as an internal and external evaluator and monitoring and evaluation specialist with multiple international organizations across Europe and the MENA region. Currently, Khalil is completing his doctoral research at the Willy Brandt School of Public Policy in Germany. His research interests include evaluation in conflict-affected and fragile states, national evaluation systems, policy evaluation, and social justice, equity-focused, and youth participatory evaluation. khalil.bitar.hoc@gmail.com.

Annika Rose Dean is currently a senior researcher at the Climate Council of Australia. She holds a Bachelor of Development Studies with first class honours, and a Ph.D., obtained in 2016, which focused on evaluating the effectiveness of climate finance in building capacity to adapt to climate change in the Republic of Kiribati. Annika was the sole lecturer of a course on research methodologies for development projects at the Australian Catholic University, which covered content on tools and approaches for evaluating development projects. She is also active in a research collaboration focused on evaluating the effectiveness of the integration of climate change and disaster risk management in community development planning processes in Tonga and Fiji Islands. annika.rose.dean@gmail.com.

Joseph (Joe) Dickman leads the Evaluation and Learning Initiative in the Climate Investment Funds (CIF), managing a large portfolio of studies and learning activities covering key topics in climate finance and drawing upon the CIF experience in clean energy, energy access, sustainable forestry and climate resilience. Prior to that, Joe was Deputy Director of Research, Evaluation and Learning at The MasterCard Foundation, and has previously led design, monitoring and evaluation activities for several non-governmental organizations, including Mercy Corps and CARE International, and served as a United States Peace Corps Volunteer in Bolivia. Joe holds a master's in public affairs from the Woodrow Wilson School at Princeton University. jdickman@worldbank.org.

Osvaldo Néstor Feinstein is a professor at the Universidad Complutense de Madrid's Master in the Evaluation of Programs and Public Policies and a member of the International Evaluation Advisory Panel of UNDP's Independent Evaluation Office and of the Editorial Board of the Journal Evaluation and Program Planning. He is the editor of the World Bank Series on Evaluation & Development, having published and edited articles and books in academic and professional journals. He was a manager and advisor at the World Bank independent evaluation department, a senior evaluator at the United Nations International

Fund for Agricultural Development (IFAD) and an evaluation consultant for several international, bilateral and national organizations. He provided technical assistance to governments in Latin American and Caribbean countries, as well as in Asia and Africa, Eastern Europe, and Spain. He has also lectured and conducted workshops at Latin American, British, Finnish, Indian, Korean, Italian and Spanish universities and research centres on evaluation, development, economics and performance measurement. Osvaldo was a professor at the master's course in development studies of the Latin American Faculty for the Social Sciences (FLACSO) and he designed the Latin American Program for the Development of Evaluation Capacities (PREVAL). ofeinstein@yahoo.com.

Valerie Gordon is an International Development Consultant with over 12 years of experience in Monitoring and Evaluation. She has over 30 years of working experience in the Caribbean, 25+ years of which have been focused on environmental management, sustainable development and in the last five years, climate change. She has worked in Africa and the Pacific, having undertaken over forty results-oriented monitoring missions in these regions for the European Union external monitoring system from 2007 to the present. Thematic areas in which she has monitored include natural resources management, agriculture, climate change, sustainable energy, civil society, human rights and gender. She has undertaken evaluations for the Caribbean Development Bank, UNDP, UNESCO and the European Commission. She served on the founding board of the Caribbean Evaluators International (CEI) as Vice President and Director of Research and Publications (2014-2018), and currently serves as President (2019-2021). gordon.valerie@gmail.com.

Abdoulaye Gounou is a senior monitoring and evaluation professional. He has been Director General of evaluation at the Prime Minister's office, and he is currently the Head of the Bureau of Public Policy Evaluation and Government Action Analysis at the Presidency, Republic of Benin. He holds a diploma of Advanced Professional Studies from Senghor University of Alexandria, Egypt, in business administration and project management, and a master's degree of Laws from the National University of the Republic of Benin. He is involved in conducting process and impact evaluation of public policies, capacity development events and collaborative projects development and implementation, such as Twende Mbele, a Programme with South Africa and Uganda and West Africa Capacity building and Impact Evaluation (WACIE), a Programme with 3ie within the West Africa Economy and Monetary Union (WAEMU) that gathers height countries in West Africa. Previously, he spent five years as a technical adviser for the Government of Benin in economic strategy and reforms and private sector development. agounou0@gmail.com.

Hur Hassnain is a design, evaluation and accountability expert. He has 15 years of experience in designing and conducting evaluations in and around fragility, conflict and violence. He has lived and worked in some of the most fragile contexts in Asia, Middle East and Africa. Hur has designed, commissioned and conducted research and evaluations for projects funded by the UN agencies, European Commission and a variety of public and private donors. Hur is an expert in experimental and quasi-experimental designs, participatory and theory-based evaluation approaches and has developed innovative tools that have been used on the face of the conflict and presented internationally. Hur is the Founder at Pakistan Evaluation Association and an elected Board Member at the International Development Evaluation Association (IDEAS) for two terms. At IDEAS, Hur convenes the Thematic Interest Group on Evaluations in Fragility, Conflict and Violence (EvalFCV-ITIG). hurhassnain@hotmail.com.

Gonzalo Hernández-Licona is Member of the Independent Group of Scientists to draft the Global Sustainable Development Report 2019 for the UN; Board Member of El Colegio de México and teaches Development Economics at ITAM in Mexico since 1991. He holds a Ph.D. in Economics from Oxford University, UK, a Masters Degree in Economics from the University of Essex, UK, and a Bachelor's Degree in Economics from the Mexico Autonomous Institute of Technology (ITAM). He was the Executive Secretary of the National Council for the Evaluation of Social Development Policy (CONEVAL); Board Commissioner of the International Initiative for Impact Evaluation (3ie), member of the Evaluation Committee of the GAVI Alliance. He was Head of Evaluation and Monitoring at the Ministry of Social Development. He was full-time Chair Professor at the Mexico Autonomous Institute of Technology (ITAM) in the Department of Economics from 1991 to 1992 and from 1996 to 2002 and acted as Headmaster of the BA in Economics in the same institution. From 1996 to 2000 he was the Academic Representative to the North American Agreement on Labour Cooperation Commission. ghlicona@gmail.com.

Viliamu Iese is a research fellow at the Pacific Centre for Environment and Sustainable Development, The University of the South Pacific. He has 20 years of experience in researching impacts of climate change on food security and the effectiveness of adaptation options on improving livelihood and resilience in Pacific Island Countries. He is also active in evaluating the effectiveness of the integration of climate change and disaster risk management in community development planning processes in Tonga and Fiji Islands. He is currently a member of a multi-partners collaboration to develop methods and evaluate effectiveness of community food production initiatives to health risks in the Pacific Islands and the Caribbean regions. Viliamu has been a reviewer of both development proposals and projects in the Pacific. ieseviliamu1@gmail.com.

Cristina Magro is the Secretary-General of the International Development Evaluation Association (IDEAS), now serving her second term. As a Senior Education Expert Consultant, Cristina worked at the Ministry of Education in Timor-Leste under contracts run by the World Bank and UNICEF, acting as an evaluator of school materials and para-didactic resources, and as an advisor for the Vice-Minister of Primary and Secondary Education and for the Minister of Education. She worked for the Brazilian Ministry of Education evaluating Universities and Faculties, and the Observatory of Education. For 26 years she was a full professor at the Federal University of Minas Gerais, as well as an associate professor, visiting scholar and researcher at academic institutions in Brazil and abroad. She holds a Ph.D. in Language Sciences, is an IPDET graduate and certified in Institutional Assessment by the Brazilian Ministry of Education and Culture. She dedicated more than 25 years of research in systems thinking and the teaching of systems thinking for non-academic audiences. She has authored several publications, either alone or with academic and professional partners. magro.cristina@gmail.com.

Bianca Montrosse-Moorhead is an Associate Professor of Evaluation at the University of Connecticut, United States, where she trains future practitioners and scholars. Her research centres on evaluation education, theory, methodology, practice, and capacity building. For more than a decade, she has directed evaluations at the local, state, and national level totalling over $10 million (USD). She currently serves as a Board Member for the American Evaluation Association, as an Editorial Board Member for the New Directions for Evaluation journal, and as Chair of EvalYouth. bianca.montrosse@gmail.com.

Silvia Salinas-Mulder is a Bolivian anthropologist, with a specialisation in Andean Studies and a master's degree in Decentralization and Public Administration. Twenty-nine years of experience as a senior consultant, researcher and evaluator in the social development sector. Recognised for her specialised systemic and creative approaches to key poverty, development and exclusion topics. Advocator and facilitator of transformative planning, management and evaluation approaches. Particularly interested in power issues, ethics, gender and intercultural relations. Between 2016 and June 2018 president of the Bolivian Monitoring and Evaluation Network (REDMEBOL). Since December 2017 Coordinator of the Latin American and Caribbean Monitoring, Evaluation and Systematization Network (ReLAC). She is also Co-founder and Strategic Director of FuturaLab, a company dedicated to boost innovation culture. ssalinasmu@hotmail.es.

Otto Navunicagi is an Early Career Researcher pursuing a Master of Science – Climate Change (MSc-CC) at the Pacific Centre for Environment and Sustainable

Development – the University of the South Pacific in Fiji, supervised by Viliamu Iese. Otto's research evaluates the Impacts of Community Food Production Initiatives (CFPIs) on Food Security Resilience in two selected communities in Fiji. Otto is passionate about research and has also had the opportunity to co-author a few research papers that are currently under review. After completing his master's Otto hopes to develop a framework or guideline aimed at improving CFPIs in order to achieve Food Security Resilience. The framework when completed can be used by Food Security Practitioners. ottonavunicagi@gmail.com.

Claudia Nicholson is a Grenadian Economist with almost 25 years of extensive experience in conducting evaluation projects in the United States, Canada, and the Caribbean. She has the unique experience of working on randomized control trials for close to 20 years, a then pioneering social policy research technique, in Canada and the United States with two well renown and award-winning social policy research organisations – MDRC in New York, United States, and Social Research and Demonstration Corporation in Ottawa, Canada. Since 2013 Claudia has been working as an independent international development consultant focused on monitoring and evaluation, and survey research in the English-speaking Caribbean. cldnicholson@gmail.com.

Nadini Persaud holds a Bachelor of Science in Accounting (First Class Honours), a Master of Science in Project Management and Evaluation (Distinction), a doctorate in evaluation, and two professional accounting designations CGA (Gold Award) and CPA. She is a lecturer at the University of the West Indies, Cave Hill Campus, Barbados and was the program coordinator for the M.Sc. Project Management and Evaluation degree from 2008-2017. She serves on a number of international organizations. Current posts include Chair — Finance Advisory Board, American Evaluation Association (AEA), Program Chair — CEBE TIG, AEA, and Advisory Board Member of the Faster Forward Fund Foundation in the United States. Her publications include papers and articles on evaluation, accounting, and research methods and a co-authored book titled *Research methodology basics: From conceptualization to write-up*. She has also presented 18 conference papers at the AEA on the importance of cost analysis in program evaluation. npersaud07@yahoo.com.

Jyotsna (Jo) Puri is the Head of the Independent Evaluation Unit of the Green Climate Fund. She is also an adjunct Associate Professor at Columbia University's School of International and Public Affairs. Her areas of work include policy impact analysis on poverty reduction, environment, agriculture, health, and climate change. She has more than 22 years of experience in policy research and development evaluation at organizations including the World Bank and the United Nations and has led

evaluation-related work for the United Nations Development Programme, UNICEF, and the Global Environment Facility. Her expertise is in applying mixed methods that use rigorous quantitative methods informed by high-quality qualitative approaches. She holds a Ph.D. and M.Sc. in resource economics and a master's in development economics. jpuri@gcfund.org.

Archi Rastogi is an Evaluation Specialist with the Independent Evaluation Office of the Green Climate Fund. Previously, he worked as a Consultant at Universalia Management Group. Archi has led or participated in more than a dozen large and small evaluations with innovative methodologies. These evaluations have spanned various fields, including research-for-development, climate change, agriculture, water, integrity, gender and justice. Archi was part of the Overall Performance Study of the Global Environment Facility, and Forward-looking Performance Review of the Green Climate Fund. Archi has worked in diverse sectors including a grassroots organization, United Nations, media, academia, and the private sector. He has a Ph.D. in Natural Resource Sciences, and a Masters in Environment Management. archirastogi@gmail.com.

Antonina Rishko-Porcescu is the Secretary of the EvalYouth Global Network, and Founder and Co-Chair of the regional chapter of Eastern Europe, Central Asia and South Caucasus. She has experience in both social research and monitoring and evaluation. She has a Ph.D. in Sociology from the Institute of Sociology of the National Academy of Sciences of Ukraine, and while there she focused on the problem of social integration of migrants, mainly Ukrainian migrants in big Polish cities. As an evaluator, Antonina has experience in local, national and regional evaluations. Currently, she serves the United Nations Volunteers as Monitoring and Evaluation Officer at the United Nations Population Fund Regional Office for Eastern Europe and Central Asia. antonina.rishko@gmail.com.

Jope Tarai is an emerging indigenous scholar based at the University of the South Pacific. He works in the School of Government, Development and International Affairs. His research interests include Pacific Regionalism, New Media, Pacific Politic, Development and Tuna Diplomacy. jope.tarai@usp.ac.fj.

Kseniya Temnenko is a Knowledge Management Officer at the Global Environment Facility Independent Evaluation Office with experience in knowledge management and evaluation. She works on innovative synthesis studies on priority topics relevant to GEF-supported projects, programs, policies, and strategies. She also develops training materials for stakeholder meetings and learning products (such as briefs, videos, and infographics) that share lessons from evaluations. Her previous work

in the Evaluation Office of the United Nations Children's Fund (UNICEF) was in the areas of knowledge management and evaluation systems strengthening. She also worked and consulted for the United Nations Development Programme (UNDP), the Eurasia Foundation, and the Institute for Sustainable Communities (ISC). Kseniya Temnenko has master's degree in Sociology from a University in Ukraine, and acertificate from the International Program on Development Evaluation Training (IPDET). She also attended the Certified Knowledge Manager Training at the Knowledge Management Institute in Washington, D.C. ktemnenko@thegef.org.

Chris Peter Teva is an early career researcher developing an evaluation framework for sustainability, effectiveness and increasing the impacts of adaptation food security projects in the Solomon Islands. pitapile@gmail.com.

Juha I. Uitto is Director of the Global Environment Facility Independent Evaluation Office (GEF-IEO). He has worked in the GEF and the United Nations Development Programme (UNDP) since 1999 as Senior Monitoring and Evaluation Specialist/Coordinator, Evaluation Advisor, and, finally, Deputy Director of UNDP's Independent Evaluation Office. Juha spent the 1990s with the United Nations University in Tokyo as coordinator of the university's environment and sustainable development research and training programs. He has also worked with the Nordic Africa Institute and the International Fund for Agricultural Development and has held visiting positions in universities in the United States and Japan. He was educated at the Universities of Helsinki and Lund and holds a Ph.D. in social and economic geography. juitto@thegef.org.

Morgan Wairiu is the Acting Director of the Pacific Centre for Environment and Sustainable Development (PaCE-SD) at the University of the South Pacific (USP). He manages the Centre's collaboration in Research and Teaching across the university schools, institutes and centres, as well as national, regional Governments and international development partners and across the NGOs, CSO and community sectors. With a career spanning more than three decades, he has extensive experience in environment and sustainable development across the Pacific Islands Countries. His current research interest is in climate change adaptation and risk management with a focus on food security. Morgan is an author of the IPCC Special Report on 1.5 degrees and the "small islands" chapter of the IPCC Working Group II Sixth Assessment Report. morgan.wairiu@usp.ac.fj.

Anna Williams is an independent evaluator, facilitator, and organizational learning expert specializing in complex systems change. She has worked with several international climate change funders as an evaluator, advisor, researcher, trainer, and team member. Since 2014, Anna has helped to build capacity for strategic learning at the Climate Investment Funds (CIF), including the launch of CIF's Evaluation and Learning Initiative and designing CIF's work on transformational change which she has been leading since 2016. Anna holds Masters's degrees in Public Policy and Environmental Management from Duke University and a bachelor's degree in environmental studies from Colorado College. alwilliams@worldbank.org.

Gabriela Pérez Yarahuán is currently Professor at CIDE, a Public Research Center in Mexico City and Director of the Center for Learning on Evaluation and Results in Latin America (CLEAR LAC). Her experience in the study and practice of evaluation spans for over 15 years. In her professional career she held senior positions in the Mexican government, as senior policy advisor in the Ministry of Social Development, the Ministry of Rural Development and Agriculture, and the Ministry of Education. From 2007 to 2009, she was the head of the evaluation unit at CONEVAL (National Council for the Evaluation of Social Policy in Mexico). She holds a PhD and MSc degrees from the University of Chicago and the University of Warwick. gabriela.perez@cide.edu

A call to transform our world!

Four years after Agenda 2030 was adopted by the global community, this call has become more urgent than ever. Evaluation will need to step up to the plate and contribute.

The Sustainable Development Goals require transformational change in our societies, our economies and our interaction with nature. "Business as usual" is not an option, as climate change is developing into a full-blown climate crisis, and social unrest and populist movements throughout the world threaten to undo the achievements of many decades. Increasingly, evaluators are required to come up with evidence on what promotes and supports transformational change. This book presents essays on how evaluators and evaluations can make this transition, from evaluation of projects, programmes and policies to evaluation of how these interventions could and should be leading to transformational change.

The contributors to this book consider approaches, methods and techniques, whether the capacities of evaluators and of the evaluation systems in which they are functioning are up to the task, and what can be done to enable evaluators and evaluation systems to move in a transformational direction. New approaches are introduced. A call for a youth revolution is included! Regional differences are discussed. Areas of work are explored, such as the role of heritage and cultural values in evaluation of conflict resolution, and evaluation of climate action and environment and development. Systems thinking and analysis are presented as necessary for transformational evaluations at the level of societies, economies and the environment.

ideas-global.org

Evaluation for transformational change: opportunities and challenges for the Sustainable Development Goals (Rob D. van den Berg, Cristina Magro, and Silvia Salinas Mulder, editors) is a publication that IDEAS has prepared for its Global Assembly in Prague, 2 to 4 October 2019. A grant from Universalia enabled this publication.

UNIVERSALIA

www.universalia.com

ISBN (paper): 978-1-9999329-2-3
ISBN (electronic): 978-1-9999329-3-0
Kindle: 978-1-9999329-4-7