

TRANSFORMATIONAL EVALUATION

FOR THE GLOBAL CRISES OF OUR TIMES

Rob D. van den Berg

Cristina Magro

Marie-Hélène Adrien

EDITORS



IDEAS

KNOWLEDGE CAPACITY DEVELOPMENT NETWORKING

TRANSFORMATIONAL EVALUATION

FOR THE GLOBAL CRISES OF OUR TIMES

TRANSFORMATIONAL EVALUATION

FOR THE GLOBAL CRISES OF OUR TIMES

EDITORS

Rob D. van den Berg

Visiting Professor, King's College London
Leidschendam, the Netherlands

Cristina Magro

International Evaluation Academy, Member of the Council
Belo Horizonte, MG, Brazil

Marie-Hélène Adrien

Universal Management Group, Senior Associate Consultant
Montreal, Canada



© 2021 International Development Evaluation Association (IDEAS), Exeter, UK
Email: ideascoordinator@gmail.com

All rights reserved.

Any views expressed in this book are those of the authors. They do not necessarily represent the views of IDEAS, the editors, the authors' institutions, financial sponsors or reviewers.

This book is distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC-ND 4.0): <https://creativecommons.org/licenses/by/4.0/legalcode>

This license permits any non-commercial use, duplication, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source. Please cite the work as follows:

Van den Berg, Rob D., Cristina Magro and Marie-Hélène Adrien (eds.). 2021. *Transformational Evaluation for the Global Crises of Our Times*. Exeter, UK: IDEAS.

ISBN (paper): 978-1-9168982-0-2

ISBN (electronic): 978-1-9999329-9-2

Assistant editor: Zuzana Vozárová

Copy-editing: Ann Shildneck

Design: Nita Congress

'Business as usual' is a thing of the past. The COVID-19 pandemic may be the latest in a series of era-defining challenges – it will not be the last. If humanity is to move forward, transformational change is desperately needed. Evaluation has a key role to play in how governments and partners support this change. By combining practical experiences with perspectives drawn from new initiatives, the IDEAS book offers invaluable insights into how evaluation can be transformed to support transformational change on the global stage.

*Indran A. Naidoo, Director of the Office of Independent Evaluation
of IFAD and Council Member of the
International Evaluation Academy*

If we want to see evaluation contributing to long desired changes towards a regenerative economy, social justice and care for nature, we must start to think about it and do it in new ways, supported by knowledge from other fields of science. IDEAS as the first global evaluation community called upon us to transform the field and gave us a bold vision in the Prague Declaration. This is an essential strategy for the Global Evaluation Agenda of the new decade. Across continents, educational systems, and historical complexities, this book presents the language evaluators should use. A mandatory read for all young evaluators.

*Weronika Felcis, Independent Evaluation Expert and Board member
of the European Evaluation Society, as well as Secretary of the
International Organization for Cooperation in Evaluation*

This is not the usual 'conference book'. It is capitalizing on the changes we all have been experiencing these past few years but that we decided to share in lessons and experiences. A book that is projecting us all to the present situation of the pandemic with this question: how are we all going to transform evaluation to get the decision-makers of the world to transform the way they use evaluation?

*Bali Andriantseho, Ing, MPDI, Board member of the African
Evaluation Association and representative of AfrEA in the
International Organization for Cooperation in Evaluation*

This is a must-read, deep, timely, thought-provoking contribution, addressing sustainable development, the environment, equity, justice and giving a voice to the legitimate wants of the under-privileged. This book offers original, visionary discourse and critical perspectives on the challenging global evaluation world in the post-COVID-19 pandemic era.

Doha Abdelhamid, Member of the Academy of Scientific Research and Technology, Egypt, Founding Member of EvalMENA, and Former Board Member of IDEAS and IOCE

This evaluation book is such a delight. Each section in this book is designed to respond to field-identified needs for specific guidance that did not appear to be available in existing publications. Being an evaluator, the lessons to take from it are powerful. Loved it!

Humayun Khan, General Secretary of the Pakistan Evaluation Association and of the Asia Pacific Evaluation Association

We can achieve transformational change if we can make evaluations transcend beyond the written word. This requires a comprehensive and committed approach to the analysis of the effects of interventions on the environment, culture, the economy and health. Consequently, the professionalization of evaluators opens up new challenges. After reading these chapters you will have a sharper look at what is relevant when managing or doing an evaluation, and you will notice that 'business as usual' will no longer be an option.

Janett Salvador, Independent evaluator, co-founder of the Mexican Evaluation Society (ACEVAL), Former Treasurer and Co-coordinator in the strengthening process of ReLAC

Dedication

Remembering Sulley Gariba, Founding President of IDEAS... and so much more...

A few months ago, our dear friend Sulley Gariba shared with some of us an eternal truth: 'Never give up on what you really want to do. The person with big dreams is more powerful than one with all the facts.' And this, he continued with the wonderful laugh that characterizes him, despite our beliefs in facts and evidence. Sulley had a cheeky mind.

And now Sulley has passed away quietly after a short illness, in his home country Ghana, leaving us, the IDEAS community with the incredible challenge to build on his legacy and to continue to bring evaluation to the service of our communities.

As a thinker and a doer Sulley would tell us to remember him for the dreams he fulfilled, amongst which the creation of IDEAS. As he and some thought leaders gathered in 2002 to create our Association, little did he anticipate the strengths of IDEAS's footprint in the evaluation community. Led by Sulley from 2002 to 2005, a core group of evaluators and policy-makers planted the seed to the current vision and mission of IDEAS, to build capacities in evaluation, to advocate for evaluation and to facilitate members' networking and engagement in evaluation activities. Sulley, we thank you all for your dedication and for your vision.

Sulley was much more than a leader in the international evaluation movement. He was also an excellent evaluation practitioner. For nearly 30 years he led and contributed to the design and implementation of systems for social policy analysis, monitoring and evaluation of poverty reduction and development effectiveness. At the policy level, he served as a senior Policy Advisor for the President of Ghana and led the development and implementation of a comprehensive development strategy for the Northern

Savannah Regions of Ghana where he was born. In most recent years, he had moved to Canada to serve as Ghana's High Commissioner to Canada (2014–2017) and, in 2019 was appointed Head of Country and Regional Program Impact for the Mastercard Foundation.

Sulley, we will miss your smartness, we will miss your insights into addressing the complexities of our world, we will miss you, a Ghanaian brother willing to lend a hand to the new generation of young evaluators.

On behalf of your IDEAS family, we, who succeeded you as President of IDEAS, wish you goodbye and pledge individually and collectively that IDEAS will continue to fight for social justice in the world and for the need to bring evidence to power.

Marie-Hélène Adrien, Ray C. Rist,
Rob D. van den Berg and Ada Ocampo

Contents

Dedication	vii
Contents	ix
Foreword	xiii
Acknowledgements	xv

Part I: Setting the Stage xvii

Chapter 1: Transforming Evaluation for Transformational Changes 1

Cristina Magro, Rob D. van den Berg and Marie-Hélène Adrien

Chapter 2: Blue Marble Evaluation Perspective: How Evaluations Help Solve Global Crises 19

Michael Quinn Patton

Chapter 3: Evaluation in an Uncertain World: A View from the Global South 37

Adeline Sibanda and Zenda Ofir

Part II: Experiences 63

Chapter 4: Ensuring Transformational Change for Climate Action . 65

Matthew Savage, Tim Larson, Jessica Kyle and Sam McPherson

Chapter 5: Evaluation at the Nexus Between Nature and Humanity for Transformational Change 87

Juha I. Uitto

Chapter 6: Contradictions and Complementarities Between South and North on Transformation in the Anthropocene105

Mehjabeen Abidi-Habib, Jane Burt, John Colvin,
Chimwemwe Msukwa and Mutizwa Mukute

Chapter 7: Governance Pathways for the Greater Caribbean: Transformative Evaluation Principles133

Lennise J. C. Baptiste

Part III: Professionalization 155

Chapter 8: To Be or Not to Be an Evaluator for Transformational Change: Perspectives from the Global South . . 157

Pablo Rodríguez-Bilella, Silvia Salinas Mulder and Sonal Zaveri

Chapter 9: Avenues for Young and Emerging Evaluators' Journeys to 2030177

Kenza Bennani, Marie-Hélène Adrien and Gerardo Sánchez-Romero

Chapter 10: Towards an International Evaluation Academy as an Agent of Transformational Change.199

Linda G. Morra Imas

Chapter 11: Evaluation for Transformation: What Will It Take? . . . 217

Robert Picciotto

Part IV: Themes and Cases. 239

Chapter 12: Multiple Dimensions of Evaluation in Fragility, Conflict and Violence241

Inga-Lill Aronsson and Hur Hassnain

Chapter 13: The Potential of Value Chain Development for Transformative Change: The Experiences of AfDB and IFAD . . . 263

Fabrizio Felloni and Girma Kumbi

Chapter 14: Evaluating Under Fragility: Lessons from the Palestinian Context 287

Khaled Rajab

Part V: Approaches and Methods 309

Chapter 15: Complex Systems, Development Trajectories and Theories of Change 311

Aaron E. Zazueta, Nima Bahramalian, Thuy Thu Le,
Johannes Dobinger and Eko Ruddy Cahyadi

Chapter 16: A Complexity-Based Meta-Theory of Change for Transformation Towards Green Energy 345

Jonathan A. Morell (Jonny)

Chapter 17: Of Portals and Paradigms: Evaluation, Systems Thinking and the Pandemic 371

Cristina Magro and Rob D. van den Berg

Part VI: The Prague Declaration 395

Chapter 18: The Prague Declaration: Meaning and Testimonials . . 397

Rob D. van den Berg, Daniel Svoboda, Ada Ocampo, Juha I. Uitto,
Silvia Salinas Mulder, Rashmi Agrawal and Josephine Watera

Contributors 435

Foreword

This book, the sixth following an IDEAS Global Assembly, is being finalized 14 months after the World Health Organization declared the COVID-19 pandemic in March 2020. At this point, although the world is still facing the challenges and consequences of this unprecedented global crisis, there seems to be a light at the end of the tunnel: vaccines are being rolled out. Although there is still a long way to go to ensure coverage, as well as equal access to vaccines, hope is emerging. At IDEAS, as Rob D. van den Berg, co-editor of this book and President of IDEAS when the book was conceptualized, says in his statement in the last section of the book, we 'hope that, while the world is slowly breaking free of the clutches of the COVID-19 pandemic, this book may function as a source for rethinking and transforming evaluation to better serve the world'.

IDEAS is proud to release a resource that smartly combines conceptual and methodological analyses of and proposals for transformational evaluation and its potential to contribute to transformative development, with a range of practical experiences and testimonies. Additionally, this book discusses the professionalization of evaluation, a topic that is crucial in the discussion of how to ensure that evaluation will remain relevant. The chapters on professionalization include a thorough discussion of the perspectives of young and emerging evaluators from various parts of the world and new initiatives on professionalization such as the International Evaluation Academy. The book closes with a special section on the Prague Declaration of 4 October 2019 that participants of the IDEAS Global Assembly and the Third International Conference on Evaluating Environment and Development endorsed, as well as statements on and practical suggestions for implementing the Declaration.

IDEAS would like to encourage the evaluation community, as well as researchers, government officials, parliamentarians and the wider public, to read this book and to comment on it. Your feedback is important to us, especially because we are committed to continuing to produce and publish books.

IDEAS dedicates this book to our first President, friend and colleague Sulley Gariba, who passed away just a few days before this book was closed, on 27 April 2021. Sulley, we remain truly grateful for your leadership, insights and friendliness. Heartfelt thanks for initiating IDEAS and contributing to its growth. We remain committed to following your path.

Ada Ocampo
IDEAS President

Acknowledgements

The editors would like to thank all authors for their contributions, and their willingness to deliver despite the tough pandemic times which have shaken our routines and minds! All authors have been identified by their full names, without titles, in their chapters, while their biographies are included in the Contributors session of the book. The bios contain further information on their background and achievements, and an email address for contact.

A special thanks to Ada Ocampo, the new President of IDEAS for the period of 2020–2023, for her contribution with the Foreword and for her statement on the Prague Declaration.

References to Internet sources have been checked and verified in April and May 2021. Formulations such as 'available at' and 'accessed on' have not been used, as Internet addresses are easily recognizable and are subject to change. They are given as a useful tool, but not as cast in stone. We hope the readers benefit from them.

We very much thank Nita Congress and Ann Shildneck for their great contribution to the wonderful design and layout of the book, as well as the very thorough and enriching copy-editing of the chapters. They have ensured a more accessible publication, visually and textually.

The amazing collection of authors here gathered, coming from all over the globe, wrote their contributions in varieties of English which, for the sake of publication, have been harmonized. IDEAS as a UK publisher and a global association follows the Oxford Spelling as is common for academic writing for an international readership and as adopted by the United Nations and many other international organizations.

The editors would like to especially thank Zuzana Vozárová, who supported us as assistant editor. She played a key role throughout the process

of initiating the book, inviting authors, interacting with authors on deadlines and on any questions they had. Her watchful eye was much appreciated on the editing process, from comments on drafts to alerting editors and authors when steps in the process were due, and supporting editorial meetings with her overviews of where we were and her lightning speed in presenting the minutes of meetings, as well as keeping our files in order. We have very much enjoyed working with you, always with a smile and with full support for what we wanted to achieve. You have played an important role in ensuring this book took shape. Many thanks and we wish you the best in your future career.

Rob D. van den Berg,
Cristina Magro and Marie-Hélène Adrien

PART I

SETTING
THE
STAGE





CHAPTER 1

Transforming Evaluation for Transformational Changes

CRISTINA MAGRO, ROB D. VAN DEN BERG AND
MARIE-HÉLÈNE ADRIEN

When this book was taking shape, the unexpected COVID-19 pandemic was creating havoc throughout the world. Its occurrence has demonstrated the enormous challenges humanity is facing. The world is becoming more unpredictable by the minute, and speculations arise as to how we will survive the pandemic, when the next one will start, and when and how we can recover. Many of us seem to be willing to recover by transforming our societies, our economies and our relationship with each other and with the environment and the planet into a truly sustainable and just one.

That business cannot continue as usual seems very clear. Our 'usual' practices and dominant human systems bring climate change, biodiversity loss, rising and structural inequality and inequity, economic exploitation and repression, increasing mobility and migration. On top of that, we experience increased insecurity and unpredictability, as perhaps most clearly illustrated by the COVID-19 pandemic.

We are less clear, though, on what we can do to move business from usual to new ways of interacting with each other and the planet on which we live. The agenda for action that almost all countries in the world have agreed upon is widely recognized as an 'aspirational' one, a shared vision

of the future that is meant to guide our actions. For its critics, it simply means 'impossible to achieve'. Furthermore, the agenda does not embody any mechanism to enforce its application, and it is powerless against lip service and wilful sabotage. It is, moreover, open ended, and the actions proposed are in themselves not sufficient for humanity to survive. On top of that, large numbers of our fellow citizens are not convinced at all that we face existential crises and vote for representatives and governments that undermine any common action.

A Call for Action without a Central Machinery

The aspirational agenda for action, as laid down and agreed upon in Agenda 2030 and its Sustainable Development Goals (SDGs) and the Paris Agreement on Climate Change, implies the decentralization of its implementation. Although effective international partnership is one of the stated aims of the SDGs, each country has the sovereign choice of how it would collaborate. It is up to each country to set its own priorities and find its own route to solve our global crises. It is thus a call for global transformational change without a centralized global mechanism to control and orchestrate global action.

The SDGs include indicators and targets, and countries are invited to report on progress towards the goals, according to the priorities they set. Although data gathering and analysis is primarily the way evidence on progress is tracked, for the first time in history, a role is ensured for evaluation. Agenda 2030 calls for 'follow-up and review processes at all levels' (UN 2015, 31–32), with evaluation tasked with providing evidence that is rigorous on top of the data required to track progress. Agenda 2030 recognizes (UN 2015) that many countries in the Global South need financial support to enable them to build up and strengthen their national evaluation programmes, which in many cases are more advanced than similar systems in the Global North but are not reaching their full potential because of low budgets. Since 2015, the SDGs and the role of evaluation have been discussed in the United Nations, in the multilateral banks, in countries and in professional evaluation circles, in which IDEAS has played a role. Exchange of experiences in evaluation between the Global South and the Global North have been stimulated in all IDEAS conferences, strengthening capacities and exploring and amplifying the perspectives to be considered by evaluators from both hemispheres. Partnerships with different kinds of institutions

were formed, promoting more intensive interaction. In 2015, in Bangkok, the IDEAS Global Assembly focused on evaluation for sustainability, asking for attention to be paid to the inadequacy of 'business as usual' and the need to change. In 2017, in Guanajuato, the Global Assembly discussed the role of evaluation in the SDGs, followed in 2019 in Prague by the focus on the role of evaluation in transformational change. This book emerges out of the Prague conference. Our writers have been challenged to build upon what was discussed in Prague and reach out towards the future of our profession.

The main problem remains that, although transformational change is desperately needed, and evaluation has contributed to better understanding of how governments and their partners can support and strengthen transformational change, we still lack a broad understanding of and agreement on what constitutes transformational change; how it can be supported, initiated and strengthened; and when and how it can be termed successful and effective in warding off disaster.

This is a somewhat curious situation, because transformational change is well known to mankind. We have lived with transformational changes throughout history, many of which are known through terms and concepts such as globalization; colonization; the agricultural, industrial and energy revolutions; commercialization; the various information, communication and technology revolutions; and an understanding of the roots of our global crises. Knowing how we happened to be where we are today does not mean that we have a clear perspective on how to do, stimulate and support potential transformations. One could argue that most if not all of the transformations of the past happened to humanity without any master plan and without global collaboration or agreement on the goals to be achieved. One may conclude, 'No wonder it is such a mess, and no wonder urgent action is needed', or sceptics could claim 'Transformations happen throughout history, and no action seems necessary or possible'.

The Need to Transform Evaluation

Evaluation is not immediately qualified to help humanity. It is not immune to 'business as usual'. Many evaluators feel that they should just continue to do their well-designed, methodologically sharp, properly implemented evaluations, often focusing on whether interventions have produced counterfactual evidence on 'what works', or whether money, time and effort were spent efficiently and effectively and learning and benefits were achieved.

Historically, evaluation has been closely linked to accountability; in international cooperation, the Global South is often required to present proof of a well-implemented project to the donor, often serving the accountability needs of the sponsor and not the learning needs of the country or the unit being evaluated. Questions about longer-term impact and sustainability tend to be put aside as 'not within the time frame of the evaluation' or briefly explored through the potential to continue to bring benefits to recipients through learning. Evaluators often argue that they hardly receive sufficient funds to answer the more challenging questions, leading to a micro-macro paradox of successfully implemented projects, with substantial outcomes that demonstrate 'micro' success, which is not reflected at the 'macro' level of countries and global challenges, as Van den Berg and Cando-Noordhuizen (2017) have demonstrated for global action on climate change. The straitjacket evaluation model that commissioners have traditionally imposed on evaluators, with its emphasis on accountability and meeting the needs of donors, shows the need to transform evaluation itself to enable it to contribute to transformational change.

Perhaps one of the most influential groups unwittingly promoting business as usual is the one that is focused on evidence-based policies. One of its core beliefs is that policies for which no rigorous and counterfactual evidence is available should not be adopted. Change should be based on experimentation and finding out whether there is hard evidence that the change 'works', but politics and policies often must be based on what is needed, rather than on hard evidence based on experiments, because this evidence will not be available in time, and may never be available. Critics of the evidence-based movement, such as Paul Cairney (2016), focus on the 'bounded rationality' of policy processes, where much more must play a role than rigorous evidence, or the lack of it. Justin Parkhurst (2017) asks for attention to be paid to 'good governance' in using evidence for policymaking, focusing on the appropriateness of evidence for the problems that must be addressed. If evaluation remains stuck in its tradition, it will have great difficulties delivering relevant material for transformational change. Evaluation must reframe its purposes, and its tools, methods and approaches must be chosen considering an interwoven network of changes to be achieved and reflecting the urgency of finding solutions and a way forward.

Within the evaluation profession, a relatively small group of evaluators and evaluation units have contributed in a major way to better understanding of how evaluation could support transformational change and be transformed itself in the process. These efforts started more than three

decades ago and had even earlier roots. They are and were scientifically linked to systems thinking. In evaluation, they were initially developed for climate change; for understanding the interactions between ecosystems and humanity and for delivering our natural infrastructure of clean air, healthy soils, healthy foods, water, energy and so on. For a long time, any evaluation approach at the nexus of development and environment was treated in mainstream evaluation land as of hardly any consequence for international evaluation practice. The first international conference on evaluating climate change and development took place in Alexandria, Egypt, in 2007. The resulting book was the first of its kind as well (Van den Berg and Feinstein 2009), incorporating a great range of authors and chapters that initiated discussions that would reverberate in later years and grow to include other environmental issues. Publications such as those of Juha Uitto (2017; 2021) and the Third International Conference on Evaluating Environment and Development, which took place jointly with the IDEAS Global Assembly in Prague in October 2019, promoted and showcased this relevant growth.

Systems thinking and approaches were also presented in a broad range of articles in the *American Journal of Evaluation* and the *Evaluation Journal*, as well as *New Directions in Evaluation*, by authors from all around the globe. It was discussed in the IDEAS publication *Evaluation for Transformational Change* (Van den Berg, Magro and Salinas Mulder 2019), including a majority of authors from the Global South, with more women than men, from all continents and from small island developing states. Contemporary scientific ways of thinking can support transformational changes (Patton 2020a; Magro and Van den Berg 2019). Extensive work in evaluation has been directed to support evaluators in promotion of the learning that transformations require through, for example, redefinition of its object (Uitto, Puri and Van den Berg 2017); redesign of evaluation criteria (Patton 2020b), instruments, values and principles embedded in the evaluation work; review of the relationships between the projects, programmes and policies and the systems changes observed (Ofir et al. 2019); and the relationship between monitoring and evaluation for adequate follow-up of the intended changes (Chaplowe and Hejnowicz 2021).

A second stream of transformational work has focused on social and democratic transformation, necessary for making our societies inclusive and leading to social justice. An example is the transformative lens that Donna Mertens (2009) advocates, which came to include many aspects that have gained ascendancy in recent times, such as Indigenous perspectives (Mertens, Cram and Chilisa 2013). The link from social justice to democratic values has also been made, as Robert Picciotto (2015) discusses in his

article in the *Evaluation Journal*. He sees a transformational role for democratic evaluators to provide neutral information services, broker debate and facilitate deliberative decision-making processes and thus describes an evaluative role supporting the shift to governance of evidence that Parkhurst (2017) advocates. This paragraph could not do justice to the many other efforts to enable evaluation to support transformational change of our societies.

A third stream of transformational thinking and science is linked to economics, unfortunately always labelled as 'heterodox' to express that it is not mainstream economics, that continues to build on neoliberal perspectives. Ha-Joon Chang (2007), Kate Raworth (2017), Jason Hickel (2020) and Mariana Mazzucato (Jacobs and Mazzucato 2016; Mazzucato 2018) are four examples of influential economists who aim to transform economics in the direction of a balanced approach with social justice, sustainable economics and environmental sustainability. Apart from the fact that mainstream economists continue to marginalize them, evaluators have yet to explore their full potential.

Nevertheless, transformation of evaluation is taking place, even if in a somewhat haphazard and uncoordinated manner. A shift is definitely occurring that is not complete and is just as aspirational and full of wishful thinking and hope as the efforts to transform our interaction with each other and with the environment, before the global crises lead to catastrophes. The commitment to a collective global transformation, as promised in Agenda 2030, demands extensive coordination of actions, especially in a context in which rising traditionalism contributes to the destruction of achievements of civilization processes and cultures (Teitelbaum 2020). When fake news, denialism and escapism have shredded the social fabric and prevented the development of meaningful conversations, mutual understanding becomes increasingly difficult. On top of those confounding factors, the COVID-19 pandemic has ravaged the globe, calling into question the aspiration for the SDGs and for a better future for all.

This discourse and the accompanying confusion are not new for evaluators and those working in development. The idea of *development* itself emerges from the perspective of abundance of the West and the North (broadly speaking) and is inattentive to the variety of ways of living by different people from various cultures inhabiting the planet. Questioning this idea has raised a warning for evaluators. Evaluators have addressed this in many ways, including by paying attention to multiple cosmologies, modes of thinking and living of Indigenous peoples (Mertens, Cram and Chilisa 2013), non-Western philosophies (Mishra 2017) and Black perspectives

(Coates 2017) that could bring light to this dark era and to their professional roles. The recourse to various non-traditional, non-canonical ways of thinking unveils the angst of evaluators to beat the traps of Western commonalities and to bring diversity of perspectives into play, a move that rejoins a communal and democratic demand and the quest for fundamental changes in our societies, economies and interactions with the environment.

The Need to Understand and Address Transformation

The word *transformation* is ordinarily used to refer to changes an observer sees as *major modifications of an object or phenomenon in a relatively short period of time* and often describes as if something had changed into something else or that, after the process identified as transformation occurred, there was a totally different object or phenomenon from before.

Change is universal, as Heraclitus said, '*panta rhei!*¹', but a river stream- ing through a valley changes at every moment while still remaining the same river through the valley.

Transformation is not new to humanity, although its treatment from a systems perspective and as a systems phenomenon is relatively recent. The Neolithic revolution, also called the agricultural revolution, is usually thought to have started about 12,000 years ago, when humanity adopted and developed agriculture and shifted from a lifestyle of hunting and gathering to one of settlement and domestication of plants and animals. Several geologists have argued that the agricultural revolution could be seen as the start of the Anthropocene (the age of mankind) because settlement and agriculture quickly spread throughout the world and may have led to the stabilization of the climate for an unprecedented 11,000 years that made agriculture so successful in producing food for growing populations (Lewis and Maslin 2018). One aspect of the agricultural revolution worth emphasizing is that actions developed in one system (human beings working the land planting and harvesting crops regularly) affected another system with which the first relates, which is the change in the composition of the atmosphere and climate stabilization. Because systems are not distinguished in

¹ 'Everything flows!' Heraclitus (535–475 BC), Roman philosopher, which Heraclitus expresses as 'No man steps into the same river twice'. As quoted in Plato, *Cratylus*, section 402a.

a void and establish ongoing relations with other systems, when talking about *transformation*, one should look for changes beyond the original system considered for our actions and observations. In the example above, change in the composition of the atmosphere and climate stabilization are of primary interest. Nevertheless, other systems have also changed following the agricultural revolution.

Development of states followed the agricultural revolution, which led to (international) trade, expansion, warfare and in the end, imperialism, with a few European countries colonizing the world while they expanded slavery into a cross-continental trade. Europe in general was also the site of the first industrial revolution, which led to increased, systematic use of fossil fuel energy sources and thus to human-induced climate change. Other countries like the United States and Japan started contributing to further industrial revolutions, and more countries, like China, became involved in the various information revolutions. These processes have led to more globalization and growing inequity and inequality between and within countries, issues that the SDGs and the Paris Agreement aim to redress, but also led to the rise of worldwide populism and a resurgence of racism and discrimination.

Simon Lewis and Mark Maslin (2018, 348–349) define historic systems from a perspective of how humanity is organized according to ‘modes of living’. The first agricultural revolution changed humanity from hunter-gatherers to agriculturalists. The next phase they describe starts after the European Middle Ages, when mercantile capitalism enabled some European countries to enter into international trade, including establishing trading posts and colonies all over the world. Industrial capitalism enabled imperialism and the division of the world between the major imperialist powers, which can be seen as a step towards globalization. Consumer capitalism pushed the system into overdrive and began to approach planetary limits – using resources beyond what could be replenished.

The same phenomena described above can be identified in the case of the smartphone market and climate change. Systems dynamics are such that the changes they perform may trigger events distant in time and space, in a non-linear, retroactive way – or to use common jargon in some systems approaches, in recursive feedback loops. According to this, no phenomenon in systems is equivalent to traditional linear cause-and-effect relationships. Moreover, leverage points to support and consolidate transformational changes may be within neighbouring systems and may be unpredictable but desirable as opportunities for change. All this alerts us that a larger picture, or a view of the ‘macro’, is permanently needed.

These revolutions are all seen as leading to transformational changes in human societies, economies and our interactions with nature and our environment. These processes continued as decolonization took shape after the Second World War and new revolutions emerged, with the information revolution and the emergence of computing power, the Internet and social media, among others. As we look at these historical processes from a systems perspective and move away from more or less static, individualized and linear event perspectives, it becomes clear how complex and powerful systems structural interrelations are. The critical thing is that these processes seem to happen without conscious guidance by humanity, and when a trajectory has become firmly established, as with fossil-fuel energy sources, it seems almost impossible to change, but change we must, and conscious guidance we must provide, or we will not survive.

What This Book Offers

This book offers an introduction to the manifold problems we encounter when we aim to establish how evaluation can be transformed to support transformational change. The first section sets the stage for the grand challenges we are facing. The second section delivers four chapters on practical experiences of evaluators with transformational evaluations. The third section raises issues of professionalization, including perspectives of young and emerging evaluators, new initiatives like the International Evaluation Academy, and deep and broad reflections on professionalization in evaluation. The fourth section goes into themes and cases, and the fifth section presents detailed approaches and methods. The book ends with a discussion of the Prague Declaration of 4 October 2019, and whether it calls on us to further develop concepts and approaches.

The first section includes this introduction, which explores the context of the book and presents historical and systemic ideas for formulating the notion of transformation. Michael Quinn Patton also sets the stage in chapter 2 on the Blue Marble Evaluation Perspective. He establishes that the global crises of our times require transformational change. For evaluation to help in achieving this, he proposes a series of principles from global thinking to knowing and facing the realities of the Anthropocene to transformative engagement and integrating these principles into the evaluation of systems change and transformation initiatives. He develops these principles as ethical in nature and as imperative for evaluators who are committed to transformational change, using the COVID-19 pandemic as illustration of

how social justice, our interactions with the environment and our economic paradigm are intertwined and need to be untangled and transformed. His chapter ends with a thorough examination of the criteria that the international evaluation community uses and how these criteria must be adjusted for evaluating transformation. This move includes a complex systems framing criterion, eco-efficient full cost accounting, adaptive sustainability with a focus on diversity, equity and inclusion, amongst others, ensuring that these criteria can contribute to a visionary, future oriented evaluation for a just, sustainable future.

In chapter 3, Adeline Sibanda and Zenda Ofir explore evaluation in an uncertain world from a Global South perspective, arguing that the global evaluation community needs to think and work in fundamentally new ways. The potential of the Global South – in which they include minority Indigenous societies around the world – must be realized, not only because global problems affect it the most, but also because it has much to offer in shaping the paths and goals of transformation towards sustainable development that the world so urgently needs. Sibanda and Ofir plead for the Global South to address the consequences of centuries of colonization, as well as ongoing power asymmetries in global governance, economic and financial systems. They see it as essential that all those who shape and work in the global evaluation system collectively ask how decolonization of the mind and of practice in the Global South and Global North can be achieved and propose conducting, synthesizing and learning from evaluations of South–South cooperation and innovative approaches to sustainable development, advancing evaluation in support of systems change and transformation and thus contributing to the development of mechanisms for intensive generation, documentation and accumulation of innovative Global South approaches to the challenges the world faces.

Chapter 4, the first of the section on experiences with transformational change evaluation, discusses the lessons learned and insights of the Climate Investment Funds (CIF). The authors, Matthew Savage, Tim Larson, Jessica Kyle and Sam McPherson, have been involved in various roles in developing a conceptual framework for the independent evaluation of transformational change in CIF investments and ongoing work to develop this framework for practitioners of climate action in multilateral banks and in countries that aim to achieve transformation. Although the findings of the evaluation played an important role in extending the CIF into a second phase and solicited a first prize in 2019 as most influential evaluation at the IFAD/World Bank/IDEAS award competition for transformational evaluations, the ongoing work is especially relevant for future practice as it aims to identify signals

of and progress towards transformational change as useful instruments for design, inception, adaptive management and evaluation of transformational initiatives.

In chapter 5, Juha Uitto tackles evaluation at the nexus between nature and humanity based on the discussions and presentations at the Third International Conference on Evaluating Environment and Development, which ran in conjunction with the IDEAS Global Assembly in Prague in October 2019. Uitto, like Michael Quinn Patton in chapter 2, uses the COVID-19 pandemic to demonstrate the interdependence of human and ecosystem health. He highlights that the global environmental crises of climate change, biodiversity loss and ecosystem degradation must be addressed through transformational changes in major systems ranging from energy and transportation to agriculture and cities. Evaluation can and should help identify solutions for the future, as the evaluative work on the portfolio of the Global Environment Facility demonstrates. According to Uitto, this requires an open theory-of-change approach that pays attention to unanticipated consequences for the environment; for different groups of people, especially the vulnerable; and for incentives and disincentives for a true transformation to sustainability.

The Emerald Network, represented by Mehjabeen Abidi-Habib, Jane Burt, John Colvin, Chimwemwe Msukwa and Mutizwa Mukute, has contributed with chapter 6 on contradictions and complementarities between the Global South and the Global North. As an emerging community of evaluation and learning with Global South and Global North consultancy partners, their experience goes beyond evaluation into policymaking, design, strategy, finance, implementation and research. They write about transformative development pathways in service to a just, regenerative, low-carbon, resilient world. They reflect on how their praxis has evolved over the past eight years, sharing stories of success and failure and what they have learned. In an innovative dialogue, they describe, explore and argue about the role of evaluation praxis in transformational design for sustainable development, focusing on how assumptions and framings, leadership teams, collaboratives and social movements seeking to address global-to-local problems such as the pandemic and climate change will be better equipped to navigate power in South–North complementarities and contradictions.

Chapter 7 brings us the perspective of the small island developing states. Based on priorities for post-COVID-19 recovery in the region that the Association of Caribbean States identified, Lennise Baptiste examines how transformative evaluation practices can help in the crisis management

and recovery phases. She argues that, considering the gaps in governance systems of the region, which were designed to keep citizens safe and provide relief in times of crisis, strengthening internal monitoring and evaluation systems would increase the accountability of member states for their responsibilities to their citizens. She identifies four pathways along which better governance and progress towards the SDGs can be achieved. She concludes that capacity building, strategic planning, policy development and the use of information and communication technology are key transformation pathways for the region.

The third section is dedicated to the professionalization of evaluation. It begins with chapter 8, in which three authors from the Global South, Pablo Rodríguez-Bilella, Silvia Salinas Mulder and Sonal Zaveri, hypothesize that, in the current neoliberal context, evaluation runs the risk of becoming another service that gives answers that those who pay for it want. Based on theory and practices from the Global South, the authors present a framework for transformative evaluation, differentiating it from conventional evaluation, and introduce a competencies profile for gender-transformative, context-relevant evaluators that unites the technical, ethical and political dimensions of transformative evaluation in a set of competencies for evaluators. The chapter identifies factors and evaluator competencies that facilitate usable evaluation and concludes by raising readers' awareness of the complexities underlying frequently invisible power issues and relations and the need to fine tune one's ability to identify and address them in evaluations.

In chapter 9, we encounter two young and emerging evaluators (YEEs), authors Kenza Bennani and Gerardo Sánchez-Romero, joining one of our editors, Marie-Hélène Adrien, in exploring avenues that YEEs can follow to become involved in transformational evaluation. They use the concept of 'professional identity work' to present a framework that defines and differentiates the various types of evaluator identities that YEEs could explore through formal employment engagement or involvement with voluntary organizations for professional evaluation and YEE networks. These should be considered 'identity workspaces' that support YEEs in discovering, understanding and shaping who they are and can become as transformational evaluators. The chapter presents inspiration and opportunities but notes that perhaps, most importantly, YEEs must be enabled to shape their futures and their contributions to transformational evaluation and in doing so contribute to the transformation of evaluation itself.

Chapter 10, by Linda Morra Imas, focuses on one of the initiatives emerging from the IDEAS Prague conferences: the International Evaluation Academy, which is on the verge of becoming fully operational. Her chapter

is therefore to some extent a snapshot of the situation in April 2021 but one that is of high importance for the future of the evaluation profession and for better linkage between science, research, policy and evaluation theory and practice. The Academy was started in a careful, interactive process including the responses to an international survey that confirmed broad-based support for the International Evaluation Academy concept, including a focus on professionalization and what it means in our time, with transforming evaluation for transformational change as one of the key areas of work. The mission, strategy, guiding principles, values, organization, thematic strategies and business model have been developed, and this chapter is our first chance to observe how it is taking shape.

Robert (Bob) Picciotto, in chapter 11, on the importance of professionalization, asks: What will it take? He positions evaluation as one of the knowledge professions, threatened like the others through rising populism, fake news and loss of confidence in the media and the 'elite' and paints this as a transformation challenge that evaluation must meet. He proposes new policy directions to transform the enabling environment of evaluation practice. Most importantly, he argues that the public good nature of evaluative evidence must be restored and that the process of turning evaluation into a private, commercial good must be reversed. To do this, the evaluation profession must break the chains of the current market-based governance model, while at the same time, evaluation must become more highly regarded and recognized for its expertise. Picciotto recommends ensuring an ethical charter, expert knowledge, proven competencies and self-management as established principles in evaluation. This is what it takes.

The fourth section, exploring themes and cases, begins with chapter 12 on evaluation in contexts affected by fragility, conflict and violence. Inga-Lill Aronsson and Hur Hassnain draw upon their experience working in these contexts and explore ideas gathered in successive interactions they maintained with evaluators while discussing the challenging experience of evaluating in such violent, rapidly changing, unpredictable environments. More specifically, the chapter is an outcome of the background research that the authors completed, a reflection of their own personal experience working in challenging environments around the world, as well as some key highlights of the rich discussions held during the full day workshop they conducted during the IDEAS Global Assembly in Prague attended by global evaluation practitioners. The chapter explores the contexts of fragility, conflict and violence that pose particular challenges for evaluation: people are vulnerable; lack trust; live in complicated and sensitive relations and lack faith in experts, outsiders, locals and government representatives alike.

The challenges range from the definition of indicators of ‘change’ or simply understanding ‘how change happens’, to data collection and analysis and associated risks and the need to rely on people on the ground to contextualize what works and what tools should be tested or avoided. They reflect and build on the recently published guide *Evaluation in Contexts of Fragility, Conflict and Violence* (Hassnain, Kelly and Somma 2021).

Chapter 13 examines the support of the African Development Bank and the International Fund for Agricultural Research for transformational change in value chains in agriculture, starting at the local and national levels but taking the international extensions of these chains into account. The authors, Fabrizio Felloni and Girma Kumbi, see mixed evidence of transformational change for smallholder farmers and rural small-scale producers because major changes in organizational culture are required at the national and international levels to institutionalize positive changes at the local level. They conclude that evaluations of local support must take the full system level of value chains into account to strengthen support for transformational change, because this will help in understanding the complexity of value chain development and the interconnectedness of a value chain system.

Evaluation of the SDGs in the unique context of Palestine as a fragile, conflict-affected country is the subject of chapter 14 by Khaled Rajab. The ongoing colonization and the deliberate policies and restrictions of the Israeli occupation mean that several SDGs and targets cannot be achieved or monitored, including those related to conservation and sustainable use of the oceans, seas and marine resources; clean water and sanitation; the environment and transboundary concerns. Within these major restrictions, the emphasis of the Palestinian government is on mobilizing stakeholders and developing partnerships to raise institutional and community awareness of the SDGs and take as many steps as possible. Rajab concludes that a tailored, adapted approach and implementation is needed to address the political, social and economic factors that affect the transformational agenda for the SDGs in fragile states under protracted occupation.

Section 5, Approaches and Methods, starts with chapter 15, Complex Systems, Development Trajectories and Theories of Change, by Aaron Zazueta, Nima Bahramalian, Thuy Thu Le, Johannes Dobinger and Eko Ruddy Cahyadi, focusing on support of the United Nations Industrial Development Organization for transformational change in quality systems in Indonesia. The authors present a case study of the application of complex adaptive systems thinking to develop a robust understanding of the dynamics of the quality system to understand the ways and extent to

which the trajectory of that system can be influenced. They explore a mix of methods that include different conceptual frameworks, analytical tools and information-gathering techniques. This approach helps identify a manageable number of conditions for intervention by embracing complexity and mapping the extent to which the key conditions interact and influence each other. It engages technical and non-technical stakeholders from inception and across the project or programme cycle (including planning, monitoring and evaluation) and is thus likely to increase effectiveness and help sustain the new trajectory by building stakeholder ownership of technically sound strategies and outcomes.

Chapter 16, by Jonny Morell, draws from complexity science to discuss transformation to a green energy future through a meta-theory that can be applied to theories of change and theories of action. Six types of behaviour of complex systems are explored: stigmergy, attractors, emergence, phase transition, self-organization and path dependence. Because complex systems behave as they do, the recommended theory of change is sparse; it has few well-defined elements or relationships among those elements. Instead of a detailed, well-ordered approach and planning for transformational change, Morell argues that *questions* need to inspire action, ranging from the characteristics of the desired state to whether that state is a 'new normal' to whether it is qualitatively different from the old normal to what the outcome chain would look like to how coordination mechanisms between all involved actors would work. These questions, formulated in more detail with reference to the six types of complex behaviour, lead to an understanding of the complexity involved in transformational change that can guide action.

In chapter 17, *Of Portals and Paradigms*, two of our editors, Cristina Magro and Rob D. van den Berg, take the reader on a journey through systems thinking. They take the pandemic as a platform for learning and exercising ways of thinking that will help evaluators focus on key questions and approaches that are adequate for current global challenges, becoming skilled in the reasoning, language and narrative of systems thinking and crossing portals of Western conventionalities. The chapter builds on the authors' previous work (Magro and Van den Berg 2019) for the exploration of basic concepts and explores the habits of a systems thinker (Benson and Marlin 2019) to gain insights and a natural flow of reasoning in systems terms. For the authors, all social, economic, environmental, cultural and cognitive contexts are favourable for evaluators to adhere to systems thinking, and more than that, the context demands transformational actions that are better understood and fostered in systems terms.

Building on the Prague Declaration

The foundation of this book was laid in Prague, in October 2019, at the IDEAS Global Assembly in conjunction with the Third International Conference on Evaluating Environment and Development. Transforming evaluation for transformational change was a recurrent theme in keynote addresses, special sessions and many presentations, with an emphasis on bringing experiences and insights from the South to the North. The conferences concluded with the adoption of the Prague Declaration, which summed up the aspirations of the participants. Many of the chapters and ideas in this volume have links to the Declaration, and we felt it would be appropriate to end the book with a special section dedicated to the Prague Declaration, presenting it along with some testimonials and statements. Although this section could have opened the book, because the Declaration was the starting point of much work since then – not least the inauguration of the International Evaluation Academy – we thought it would be good to end with the Declaration as a stepping stone to inspire further transformational work and to expand our horizons. To solve the global crises of our times, we need evaluation to step up to the plate and contribute its share.

References

- Benson, Tracy, and Sheri Marlin. 2019. *The habit-forming guide to become a systems thinker*, 2nd ed. Pittsburgh: Waters Center for Systems Thinking.
- Cairney, Paul 2015. *The politics of evidence-based policy making*. London: Palgrave Macmillan.
- Chang, Ha-Joon. 2008. *Bad Samaritans: The guilty secrets of rich nations and the threat to global prosperity*. London: Random House.
- Chaplowe, Scott, and Adam Hejnowicz. 2021. Evaluating outside the box: Evaluation's transformational potential. *Social Innovations Journal* 5: 4 March. <https://tinyurl.com/7mbr74>
- Coates, Ta-Nehisi 2017. *We were eight years in power. An American tragedy*. New York: One World Publishing.
- Hassnain, Hur, Lauren Kelly and Simona Somma (eds.) 2021. *Evaluation in contexts of fragility, conflict and violence: Guidance from global evaluation practitioners*. Exeter, UK: IDEAS.
- Hickel, Jason 2020. *Less is more. How degrowth will save the world*. London: William Heineman.
- Jacobs, Michael, and Mariana Mazzucato (eds.) 2016. *Re-thinking capitalism: Economics and policy for sustainable and inclusive growth*. Chichester, West Sussex: Wiley-Blackwell.

- Lewis, Simon L., and Mark A. Maslin 2018. *The human planet: How we created the Anthropocene*. Pelican.
- Magro, Cristina, and Rob D. van den Berg. 2019. Systems evaluations for transformational change: challenges and opportunities. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder, 131–152. Exeter, UK: IDEAS.
- Mazzucato, Mariana 2018. *The value of everything. Making and taking in the global economy*. Penguin.
- Mertens, D.M. 2009. *Transformative research and evaluation*. New York: Guilford Press.
- Mertens, D.M., F. Cram and B. Chilisa (eds.) 2013. *Indigenous pathways into social research*. Walnut Hills, CA: Left Coast Press.
- Mishra, Pankaj. 2017. *Age of anger. A history of the present*. Penguin.
- Ofir, Zenda, Gerald Singh, Emilie Beauchamp et al. 2019. From monitoring goals to systems-informed evaluation. Insights from SDG14. IIED Briefing Paper 11. <https://tinyurl.com/yy3gmmbe>
- Parkhurst, Justin. 2017. *The politics of evidence: From evidence-based policy to the good governance of evidence*. Oxon, UK: Routledge.
- Patton, Michael Quinn. 2020a. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Patton, Michael Quinn. 2020b. Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation* 39 (4): 527–542. <https://doi.org/10.1177/1098214020933689>
- Picciotto, Robert 2015. Democratic evaluation for the 21st century. *Evaluation* 21 (2): 150–166.
- Raworth, Kate. 2017. *Doughnut economics. Seven ways to think like a 21st-century economist*. London: Random House Business Books.
- Teitelbaum, Benjamin R. 2020. *War for eternity: Inside Bannon's far-right circle of global power brokers*. London: Penguin.
- Uitto, Juha I. (ed.) 2017. *Evaluating environment in international development*. London: Earthscan from Routledge.
- Uitto, Juha I. (ed.) 2021. *Evaluating environment in international development*, 2nd ed. London: Earthscan from Routledge.
- Uitto, Juha I., Jyotsna Puri and Rob D. van den Berg (eds.) 2017. *Evaluating climate change action for sustainable development*. Springer Open. <https://tinyurl.com/y5eh3e3c>
- UN (United Nations) 2015. Transforming our world: the 2030 Agenda for Sustainable Development. Resolution A/Res/70/1 adopted by the General Assembly on 25 September 2015. New York, United Nations.
- Van den Berg, Rob D., and Osvaldo Feinstein (eds.) 2009. *Evaluating climate change and development*. New Brunswick, Transaction Publishers. World Bank Series on Development, Volume 8.

- Van den Berg, Rob D., and Lee Cando-Noordhuizen. 2017. Action on climate change: What does it mean, 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, 37–52. Cham, Switzerland: Springer Open.
- Van den Berg, Rob D., Cristina Magro and Silvia Salinas Mulder (eds.). 2019. *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*. Exeter, UK: IDEAS.



CHAPTER 2

Blue Marble Evaluation Perspective: How Evaluations Help Solve Global Crises

MICHAEL QUINN PATTON

Abstract. Given the global climate emergency and related threats to a just and sustainable world, systems transformation is the clarion call of our times. Evaluators enter the fray to assess the fidelity and impacts of hypothesized transformational initiatives and trajectories. Doing so requires solid ethical grounding. The ethics of transformation involves the interconnection between personal ethics (transforming our own behaviours), professional ethics (actively advocating a transformational stance among professional evaluators), society (examining evaluation's role in support of the public good and democratic processes) and the world (ensuring attention to and engagement with the global emergency by incorporating transformational criteria of equity and sustainability into all evaluations). This chapter examines the implications of transformative ethics for evaluation theory, practice and methods, concluding with Blue Marble Evaluation as a principles-focused approach to evaluating global systems transformation.

Blue Marble Evaluation

Blue Marble refers to the iconic image of the Earth from space without borders or boundaries, a whole-Earth perspective. We humans are using our planet's resources, and polluting and warming it, in unsustainable ways. Many people, organizations and networks are working to ensure that the future is more sustainable and equitable. Blue Marble evaluators enter the fray by helping design such efforts, provide ongoing feedback for adaptation and enhanced impact and examine the long-term effectiveness of such interventions and initiatives. Incorporating the Blue Marble perspective means looking beyond nation-state boundaries and across sector and issue silos to connect the global with the local, the human with the ecological. Blue Marble Evaluation brings evaluative thinking and methods to support those trying to bring about global systems transformation.

Blue Marble Evaluation integrates design, implementation and evaluation. Evaluators bring their knowledge and expertise to bear in the design of resilient, sustainability-oriented interventions and initiatives. When an intervention and, correspondingly, an evaluation fail to incorporate an ecological sustainability perspective, both are engaging from a closed system mindset, disconnected from larger patterns and realities – like turning a crank that is not connected to anything. It is essential for planners, implementers and evaluators at the beginning of their work together to analyse the sustainability and equity challenges that the formulation of the intervention and the implications for evaluation present. Blue Marble Evaluation premises and principles provide a framework for that initial review, ongoing development and adaptation, and long-term evaluation of systems transformation contributions and impacts (Patton 2019a; 2020a).

Blue Marble Evaluation looks backward (what has been) to inform the future (what might be) based on the present trajectory (what is happening now). Evaluators examine what has worked and not worked in the past, not just to capture history, but also to inform the future. Forecasts for the future of humanity run the gamut from doom-and-gloom to utopia. Evaluation as a transdisciplinary, global profession has much to offer in navigating the risks and opportunities that arise as global change initiatives and interventions are designed and undertaken to ensure a more sustainable and equitable future.

Global Pandemic Applications

The coronavirus pandemic has provided a glimpse into the magnitude of changes that a global emergency has set in motion. United Nations Secretary-General António Guterres (2019; 2020), among many others, has warned consistently throughout the pandemic that climate change looms over the world as a larger, farther-reaching global emergency for which COVID-19 has been but a dress rehearsal, an early warning of what lies ahead at greater magnitude, albeit slower manifestation.

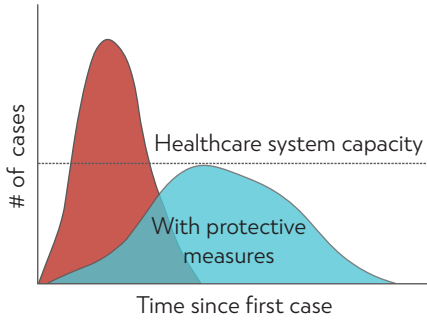
Evaluation responses to the pandemic were widespread and immediate but largely ad hoc and reactive (IEG 2020; Patton 2020b; Tolley 2020). Chelsky and Kelly (2020) of the World Bank described monitoring and evaluation during the pandemic as 'bowling in the dark'. *Better Evaluation* (2020) offered systematic, comprehensive guidance for adapting evaluation's response to COVID-19 based on the dimensions of the Rainbow Framework for Evaluation. All evaluation association conferences planned for 2020 had to be cancelled, but all associations issued statements about the continuing importance of evaluation and support for evaluators, and much evaluation training was moved online. Evaluators have been reflecting and blogging furiously and thoughtfully about what the pandemic and the climate emergency mean for evaluation (e.g. Bitar 2020; Chaplowe 2020; Feinstein 2020a; b; IEG 2020; Ofir 2020; Patton 2019b; Ramalingam et al. 2020; Vidueira 2020). Efforts abound at drawing lessons from the pandemic to inform the response to climate change (e.g. Euber 2020; Karalisi 2020).

Fundamental prevention and mitigation principles flowing from epidemiology and evaluation still apply, ignored though they may be by contemporary politicians (Mukherjee 2020). For example, the Centers for Disease Control and Prevention (CDC) *Field Epidemiology Manual*, developed scientifically over the course of decades, provides detailed protocols for addressing all aspects of a pandemic, including communications with the public. Politicians in the United States largely ignored that knowledge and wisdom because the CDC was muzzled for most of the first year of the pandemic (Duhigg 2020).

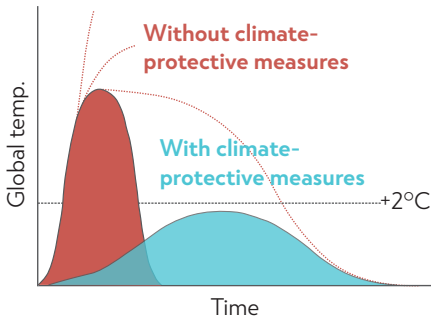
The global pandemic has provided substantial evidence to reinforce and highlight the urgency of the premise that major systems transformations are needed to address the global emergency that climate change and related global trends have brought on. Global warming; water, land and air pollution; biodiversity loss; species extinction and virulent infectious diseases pose existential threats to the future of humanity (Kolbert

Figure 2.1 Flattening the Coronavirus and Climate Change Curves

a. Coronavirus curve



b. Climate change curve



Source: a. Adapted from CDC/*The Economist* (original CC BY 2.0 E. Kim and C. T. Bergstrom); b. Hayes (2020).

2020; UNEP 2019). *The Economist* featured a cartoon showing two boxers fighting, one with the head of the world and the other with the head of the coronavirus. Observing the fight from outside the ring, but looming menacingly over it, was a much larger boxer with a fiery head wearing trunks labelled 'Climate Change'. The widely communicated and highly effective graphic that the CDC created depicting the need to 'flatten the curve' to fight the coronavirus (figure 2.1a) has been redrawn to communicate the urgent need to flatten the curve of global warming (figure 2.1b).

The pandemic has been global in scale and universal in impact, as is the climate emergency. The global climate emergency affects all of us, leading to calls for action in whatever niche we inhabit. For evaluators, that niche is evaluation, which has emerged as critically important in realizing the vision and aspirations of the Sustainable Development Goals (Rugg 2015; 2016). In this chapter,

I will illustrate the relevance of principles-focused evaluation generally (Patton 2018) and Blue Marble Evaluation principles specifically (Patton 2020a) by examining the relevance of the principles to global challenges, including the COVID-19 pandemic.

Global Thinking Principle: Applying Whole-Earth, Big-Picture Thinking to All Aspects of Systems Change

The first principle of Blue Marble Evaluation is to think globally. Certainly, the coronavirus has been a global phenomenon of epic proportions. As this is being written, more than 5 million people have been infected globally, and more than 350,000 have died. The coronavirus originated in China

in December 2019 and spread rapidly throughout the world. International agencies, governments, multinational corporations, non-governmental organizations and communities have all been directly or indirectly affected.

Inadequate collaboration among countries, lack of transparency and integrity in reporting infections and deaths and failure to follow World Health Organization and CDC guidance exacerbated the pandemic. A great deal of time was lost in January and February 2020 as many public officials, especially President Trump, downplayed the scope, scale and significance of the pandemic. Previous infectious disease threats such as severe acute respiratory syndrome, Middle East respiratory syndrome, H1N1 influenza and Ebola had been contained. The ultimate long-term effects of the pandemic and its transformative dimensions are still unfolding, but as I write this in June 2020, there is a growing consensus that there will be no return to normal. COVID-19 is proving transformative even though much of the response to the pandemic has been an attempt to contain its systems-altering significance. A major evaluation challenge will be to track, document and extract lessons from just how transformative the coronavirus turns out to be. Blue Marble Evaluation principles can help guide the search for and validation of lessons, especially global thinking and action lessons relevant not just to pandemics, but also to the looming global climate emergency.

A team of internationally recognized experts, including Nobel prize winner Joseph Stiglitz and well-known climate economist Nicholas Stern, came together to assess the economic and climate impact of taking a green route out of the pandemic crisis. They catalogued more than 700 stimulus policies into 25 broad groups and conducted a global survey of 231 experts from 53 countries, including senior officials from finance ministries and central banks. Their analysis of whether COVID-19 fiscal recovery packages will accelerate or retard progress on climate change portrays the interconnection between the coronavirus pandemic, economic policies and environmental consequences, which taken together, portray the transformations necessary to attain a more sustainable and equitable future (Hepburn et al. 2020).

The Anthropocene Principle: Knowing and Facing the Realities of the Anthropocene and Act Accordingly

The Anthropocene, according to nomenclature designation, is a new era in the history of the Earth when human beings are affecting the Earth more than natural processes are. One of those impacts is the increasing encroachment of human beings on nature. This includes more interactions

between animals and humans and between domesticated and wild animals, which have exacerbated the possibility and reality of diseases moving from animals to humans. That appears to have been the case with the coronavirus, which it is suspected came from bats in wet markets in China. As of this writing, epidemiologists are forecasting that the coronavirus pandemic will not prove to be 'the big one' but that an even greater and more devastating virus is probable.

Human exploitation of natural resources in service to prosperity and efficiency increases the likelihood of global disasters. Major shortages of protective gear, ventilators, hospital beds, pharmaceuticals and, especially in the United States, testing have exacerbated the pandemic's impact. Siddhartha Mukherjee (2020, 30), an oncologist and author of the best-selling book about the history of cancer, *The Emperor of All Maladies*, has provided an in-depth analysis of how the efficiency mania in health care administration has cost thousands of lives during the pandemic. He quotes an operations expert at Harvard Business School on the culture of efficiency:

We've been teaching how to squeeze...squeeze more efficiency, squeeze cost, squeeze more products at the same cost, squeeze out storage costs, squeeze out inventory. We really need to educate about the value of slack.

Mukherjee asks: To what extent did the market-driven, efficiency-obsessed culture of hospital administration contribute to the crisis?

His answer: The numbers in the bean counter's ledger are now body counts in a morgue.

By April 2020, when more than 4 million people had been infected worldwide, and 284,000 COVID-19 deaths had been documented, the debate shifted from public health approaches to the effects of economic depression. Trying to assess the full costs of the pandemic, a calculation that will go on for some time, should include *true-cost calculations* in which real costs to the natural environment, human welfare, equity and sustainability are included. The post-coronavirus economy will become, unintentionally, a transformed economy, 'reshaping every aspect of business' (Fortune 2020), with the nature and extent of the transformation still unfolding. The Blue Marble Evaluation Anthropocene principle guides the calculation of the costs of global actions on humans and the environment, on equity and sustainability. An example of how this can be done is the true-cost accounting framework, which measures the costs and benefits of interventions for ecosystems in ways that include human and environmental health (TEEB 2018).

Transformative Engagement Principle: Engagement and Evaluation Consistent with the Magnitude, Direction and Speed of Transformations Needed and Envisioned

The third overarching Blue Marble Evaluation principle (Patton 2020a) focuses on transformational engagement. Transformation cuts across sectors and issues. As noted earlier, the coronavirus pandemic is being described as a dress rehearsal for the global climate emergency. Transformation involves multiple, interdependent dimensions of sustainability. Caroline Heider (2017), former World Bank Group Director General Evaluation, articulated this perspective well in her reflections on the Development Assistance Committee criteria.

Taken together, these dimensions of sustainability – economic, fiscal, environmental, social – are complex. It will be difficult and costly to try to address them systematically in all evaluations. At the same time, we evaluators cannot afford to turn up with empty hands and concerns about missing data. We need to debate how we would evaluate interventions through these lenses of sustainability, see that the right questions are asked during the design of interventions and encourage the collection of relevant data (Heider 2017).

Although many are hopeful of a return to a pre-pandemic normal, the scope, depth and extent of the pandemic suggest that there will be no return to normal, nor will there be a new normal; turbulence, uncertainty and an atmosphere and reality of emergency is what the future looks like. Not only must health systems undergo transformation, but economic relationships, political institutions and societal decision-making processes are also all subject to transformation. The relationship between science and political institutions is in special need of transformation. The pandemic gave rise to the call to follow the science, but scientific protocols were not followed or not followed quickly enough, leading to the loss of thousands of lives and exacerbating the economic depression that has followed.

Nor is following the science a simple matter of following a recipe. Scientific knowledge about the coronavirus was emergent, sometimes contested among scientists, and part of the transformation needed is an infusion of evaluative thinking into political decision-making. Scientific information must be interpreted. Evaluation's special niche is synthesizing facts with values to make judgments and to do so systematically and transparently, which is the transformation needed to address future emergencies, including the global climate emergency.

Integration Principle: Integrating Blue Marble Principles into the Design of Engagement with and Evaluation of Systems Change and Transformation Initiatives

The coronavirus pandemic illustrates the importance of integrating design, execution and evaluation into a mutually reinforcing cycle that includes ongoing situation analysis, needs assessment, adaptive management and developmental evaluation. Under emergency conditions globally, there is not time to go through a sequence of in-depth situation analysis, comprehensive needs assessment, planning, design, implementation and evaluation. These things must happen simultaneously, interactively, dynamically and iteratively.

Among many other things, the global pandemic has powerfully demonstrated the interconnections among health care, school, community, economic and financial, entertainment and political systems. At any given moment, the focus has tended to be on some discrete and particular solution such as wearing masks, social distancing, more testing, quarantining the sick and flattening the curve, but the entire health care system was in crisis, an emergency that emerged and rapidly accelerated from years of neglect, ignored warnings and under-resourced health care systems at all levels. A major debate, still ongoing, is whether the problem will be solved with a vaccine or major transformations of health care systems to prepare for future pandemics and related global climate emergencies. The transformational engagement principle directs us to examine whether the actions proposed and implemented, such as giving people in the United States \$1,200, transform systems or merely treat symptoms.

The pandemic epitomizes what it means to operate scientifically and evaluatively in a complex, dynamic systems emergency. Consider the nature of epidemiology and what evaluators can learn from that esteemed and crucial profession.

Epidemiology is a science of possibilities and persuasion, not of certainties or hard proof. 'Being approximately right most of the time is better than being precisely right occasionally', the Scottish epidemiologist John Cowden (2010) wrote: 'You can only be sure when to act in retrospect.'

Epidemiologists must persuade people to upend their lives – forgo travel and socializing, submit themselves to blood draws and immunizations – even when there is scant evidence that they are directly at risk.

Epidemiologists also must learn how to maintain their persuasiveness even as their advice shifts. The projections that public health professionals make at the beginning of an emergency – for example, there is no need to

wear masks; children cannot become seriously ill – often change as hypotheses are disproved, new experiments are conducted and a virus mutates (Duhigg 2020).

Evaluators have much to learn from epidemiologists about how to engage in complex, dynamic systems during emergencies, which is the world we are all likely to face with the worsening global emergency.

Evaluation Criteria for the Anthropocene

I have offered four overarching Blue Marble Evaluation Principles: global thinking; knowing and facing the realities of the Anthropocene; transformative engagement; and integrating Blue Marble principles into the design of, engagement with, and evaluation of systems change and transformation initiatives. These principles lead to a need for criteria appropriate for evaluating systems transformation aligned with the premise that major systems transformations are needed to address the global emergency that climate change and related global trends have brought on.

Evaluation has historically focused on project and programme effectiveness. The most influential and widely used criteria for evaluating development interventions are those that the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) have adopted and disseminated (OECD DAC Network on Development Evaluation 2019). They call for interventions to be judged according to the following criteria:

- *Relevance*: Is the intervention doing the right things?
- *Coherence*: How well does the intervention fit?
- *Effectiveness*: Is the intervention achieving its objectives?
- *Impact*: What difference does the intervention make?
- *Sustainability*: Will the benefits last?

These criteria, originally formulated in 1991 and reaffirmed in 2019 (with the addition of *coherence* to the original five) are useful for those who are engaged in evaluating projects and programmes in familiar, comfortable, well-known and well-travelled ways. The 2019 revision amounted to some fine-tuning and tweaking but sent the message that the business of development and evaluation can go on as usual. The OECD website makes it clear that the revised criteria are less than optimal for addressing transformation. I have developed the ethical perspectives and challenges of the

new criteria further in my contribution to my forthcoming book *Ethics for Evaluation* (Patton in press) and discussed the criteria in depth in an article for the *American Journal of Evaluation* (Patton 2021).

The alternative to forcing the new wine of transformation into the old bottles of the DAC criteria is to bring attention to transformation by developing criteria that highlight the nature, scope and breadth of changes that the term *transformation* connotes. Responding to the systemic threats of the pandemic and climate emergency requires all hands on deck. Emergency responses, by definition, disrupt business-as-usual mindsets, modalities and methods, yet policymakers have yet to grasp the nettle, and evaluators had been mostly going about their evaluations in a business-as-usual mode, at least until the pandemic ended the pretence that 'normal' was a viable future and pushed the whole world into uncertainty about what the future holds. We now live and work in a *business as unusual world*, a post-normal world, a global emergency world, a time-is-running-out world. I therefore offer examples of alternative criteria to suggest what transformation-specific criteria might constitute and communicate. The criteria offered here (and in more detail in Patton 2021) result from two years of reflection, consultation and workshopping about criteria for transformation with others and receiving feedback. In sharing them here, I mean for them to illustrate possibilities and stimulate further contextual adaptation not to be treated as universal, standardized or mandated criteria.

Evaluation Criteria for Evaluating Transformation

Transformation Fidelity Criterion

Determine the extent to which the realities of transformational change initiatives match transformational aspirations and rhetoric. Ensure that what is called transformation constitutes transformation. Evaluate whether and how what is called transformational engagement constitutes a trajectory towards transformation.

Evaluation has a long history with the criterion of *fidelity*. There is much hype around transformation as the term has become widely used, and it has taken on a trendy cachet. Claims of transformation abound. Ensuring that such claims are meaningful and consistent with the face validity of the construct becomes a transformational evaluation priority under this criterion. Thus, the fidelity criterion aims to bring some rigour to the very notion of transformation.

Complex Systems Framing Criterion

Assess systems transformation using systems thinking principles and complexity concepts. Ensure that transforming systems is the transformational focus. Apply complex systems understandings, concepts and frameworks in evaluating transformation.

Transformation is not a project or programme. Transformational initiatives are not targeted to achieving SMART goals (Specific (simple, sensible, significant), Measurable (meaningful, motivating), Achievable (agreed, attainable), Relevant (reasonable, realistic and resourced, results-based), Time bound (time based, time limited, time and cost limited, timely, time-sensitive), which is the traditional criterion of effectiveness. Transformation means changing systems, which means addressing complexity dynamics in a world characterized by turbulence, uncertainty, unpredictability and uncontrollability. The focus of evaluation, the *evaluand* in our jargon, is *transformed systems*. Complexity rules.

Eco-Efficient Full Cost Accounting

Document and assess the full costs and benefits of systems transformations, including economic, social and environmental dimensions. Compare the full costs and benefits of baseline and transformed systems. Evaluate whether, how and to what extent transformational engagement generates net eco-efficient benefits.

Eco-efficiency offers a framework for examining transformation from unsustainable development to sustainability. This means looking beyond the traditional DAC efficiency criterion of examining the comparative costs (inputs) and benefits (outcomes) of an intervention within the boundaries of the intervention, essentially a closed-system analysis. Eco-efficiency opens and expands the analysis to examine the effects of creating goods and offering services on the use of environmental resources; effects on ecosystems; possible contributions to climate change, waste and pollution; and effects on human health, community well-being, cultural vitality and the full range of impacts on socio-ecological landscapes where humans and nature intersect. This is true-and-full cost accounting.

Adaptive Sustainability

Evaluate transformational sustainability as manifesting ecosystem resilience and adaptability at the nexus of humans and the environment. Employ a dynamic view of sustainability. Make ecosystem viability and resilience – not

programme, project or intervention continuity – the focus of sustainability. The DAC sustainability criterion focus on continuity is linear, mechanistic and static in formulation and evaluation. In contrast, at the conclusion of the 2019 IDEAs conference, participants from around the world adopted the Declaration on Evaluation for Transformational Change that included a *focus on ecosystem sustainability*.

In all our evaluations, we commit to evaluating for social, environmental and economic sustainability and transformation, including by assessing contextual factors and systemic changes. We commit to assessing and highlighting, in all evaluations, unintended negative social, economic and environmental effects (IDEAS 2019).

Diversity, equity and inclusion

Evaluate how transformational engagement manifests the values of diversity, equity and inclusion. Evaluate whether, how and to what extent transformational engagement enhances systems-level diversity, equity and inclusion. This is consistent with Agenda 2023 and leaving no one behind.

Interconnectedness momentum

Identify, understand and evaluate the interconnections that are essential and integral to transformation. Evaluate whether, how and to what extent interconnections among people, networks, institutions, ideas and movements are deepened and enhanced to support, nurture, catalyse and accelerate transformational trajectories. Evaluate whether, how and to what extent dysfunctional and constraining interconnections are disrupted and broken to liberate positive transformational energy and momentum.

These six illustrative criteria constitute an interconnected set, but in closing, let me emphasize and spotlight two of the criteria.

Equity and Sustainability

On 25 May 2020, a white police officer murdered an unarmed African-American man, George Floyd, on the street in Minneapolis, Minnesota, arresting him and keeping him down with a knee on his neck. Pleas of 'I can't breathe' were ignored. Protests and demonstrations for justice spread around the world, many turning violent, with fires, looting and additional loss of life. The economic costs will be in the billions, but the historic and current costs of racism far exceed the short-term costs of violent protests.

The costs of systemic racism to communities of colour are incalculable. For a sense of the effects of racism on people and communities, Ta-Nehisi Coates's (2015) *Between the World and Me* is a good place to start.

The Floyd incident reflected and spotlighted systemic racism, which has become increasingly visible in the pandemic. In the United States, African Americans have twice the COVID-19 infection and death rates of whites. The same disparity of infection rates and mortality between white and non-white populations shows up in countries with majority-white populations worldwide. Job losses and food insecurity as a result of the pandemic have disproportionately affected African Americans and people of colour. The anger, grief and fear resulting from the health disparities of the pandemic were a powder keg that the George Floyd murder ignited and led to social justice protests around the world.

Climate change will affect the poor and powerless substantially more than those who are economically better off and financially privileged. The world is already seeing record numbers of displaced persons and refugees. Social unrest, health pandemics, economic turbulence, political distress and ineffectiveness, societal inequities and environmental unsustainability are all linked. Each of these feeds the others. Transformational solutions must likewise be interconnected. Blue Marble Evaluation principles guide evaluation of those interconnections.

Visionary Evaluation for a Just, Sustainable Future

So how does evaluation contribute to equity and sustainability?

Policy makers and funders ask: Does the intervention work? That is an overly simple question. The more nuanced question is: What works for whom, in what ways, under what conditions and with what results? That more-complex evaluation question recognizes that no intervention works the same for everyone. Some benefit more; some benefit less. Asking evaluative questions about different effects and disparities helps address lack of equity.

The coronavirus poses the challenge of invisibility, with many people spreading the virus themselves asymptomatic, making it difficult to detect who is infected. Hatred, white supremacy and xenophobia are also difficult to detect and the effectiveness of interventions difficult to assess. Interventions aimed at the invisible and evaluations of those interventions must be aimed at general observable behaviours (wearing masks, social distancing, equal access to health care, universal access to vaccines).

Questions of sustainability become increasingly important as the pandemic rages and the global climate emergency looms. Evaluation affects these issues by making the criteria of equity and sustainability a matter of priority for *all* interventions and evaluations. We say that *what gets measured gets done*, so unless equity and sustainability are measured, those issues will not be addressed. National and international organizations and initiatives must address equity and sustainability as core criteria for evaluation at every level from the local to the global, for all projects and programmes. Making equity and sustainability universal evaluation criteria means taking them seriously, tracking them over time, making comparisons, generating findings and drawing lessons to inform future initiatives. This is how evaluation contributes to a more just, sustainable world.

Evaluation and Futuring

Futurists work on scenarios for the future; evaluators tend to be historians. When we complete an evaluation, when we have produced a report, it is automatically and instantly history, it describes and communicates what has already happened, but futurists and evaluators share the same purpose – making the future more equitable and sustainable (Patton 2019c). Evaluators study the past in hope of extracting lessons and wisdom to affect the future. Futurists run scenarios and think about possibilities and trends to affect the present. What we are both trying to do, evaluators and futurists, is affect the present so that we affect the future. When an evaluation includes recommendations, the evaluator has transitioned from being a historian, describing what has already happened, to being a futurist. Recommendations require making assumptions and having some kind of framework for thinking about the future.

What futurists have learned is that the least likely scenario is a straight-line projection of the past, yet historically, most evaluation recommendations have assumed a straight-line projection of things in the future as they have been in the past. What we are learning with Blue Marble Evaluation is to think about the future in terms of complex, dynamic systems. The turbulence, uncertainties, non-linearities and emergent phenomena in the world require ongoing, real-time monitoring. The increasing intersection between humans and nature that I discussed at the beginning of the chapter reinforces the notion that the least likely future scenario is a straight-line projection of the past. Conceptualizing and modelling multiple, diverse possible trajectories integrates knowledge of past patterns into

thinking about alternative futures. *Trajectory analysis* makes evaluations dynamic rather than static.

Trajectory analysis includes examining the likely life cycle of an intervention. Intervention effects can occur over different periods of time and evolve according to different trajectories. Although some interventions produce steadily increasing outcomes over the project lifetime, in other cases, effects may reach a maximum and then gradually decline (Bamberger, Vaessen and Raimondo 2016). Trajectory analysis to evaluate resilient sustainability requires follow-up beyond the implementation phase of a project to find out how the intervention unfolds over the long term. Evaluation designs rarely include funding for such follow-up. The evaluation ends when the intervention funding ends because funding agencies require and support evaluation only for the project intervention time period they support and for which they are accountable. For example, the project implementation phase ends when the targeted schools have been constructed, the road or irrigation system has become operational or the training programme model has been finalized, but an end-of-project evaluation is too early to assess whether sufficient capacity for resilient sustainability has been built. Blue Marble Evaluation designs should propose and incorporate resilient sustainability evaluation criteria and follow-up designs if they are to evaluate the longer-term impacts of interventions and initiatives.

Creating alternative future scenarios has become standard practice in climate change modelling. What happens with a 1.5°C increase in temperature? A 2.0°C increase? The flattening-the-curve graphic for rising global temperatures (figure 2.1b) compares alternative futures, as did the original pandemic flattening-the-curve graphic (figure 2.1a). For the 5th Assessment Report of the Intergovernmental Panel for Climate Change, four scenarios – representative emission pathways – were modelled using alternative levels of radiative forcing (global energy imbalances) and greenhouse gas concentrations by the end of the 21st century. UNICEF used the Intergovernmental Panel for Climate Change trajectory projections to create alternative scenarios of the effects of climate change on the world's children by 2050. All the scenarios showed disastrous effects on children.

Beyond climate change modelling, futurist thinking undergirds trajectory estimates and alternative scenario modelling in economics (comparing growth models), meteorology (increasingly severe weather predictions), environmental impact assessments (pollution projections), mapping the health of ecosystems under various biodiversity loss scenarios, future global pandemics, acidifying oceans and massive increases in displaced persons and refugees worldwide.

Evaluators, futurists and alternative scenario modellers in diverse disciplines can work together to create flexible adaptive mindsets. We all must figure out what data to track, what trends to pay attention to and how to separate the signal (what is important) from the noise (what is insignificant). Together, evaluators, futurists and modellers can combine evaluative thinking with strategic thinking, critical thinking with creative thinking and inferential thinking with generative thinking. That is the way forward.

Conclusion

We must all work together on global issues – the global climate emergency and related challenges – because we are all in this together. Across disciplines, across countries, across nationalities, across positions from macro to micro, across private sector and public sector, across non-governmental organizations and governments – we are all in this together. We must combine our energies, knowledge and thinking to address these major, long-term, global Blue Marble challenges. Blue Marble Evaluation is a part of entering into transformational engagement for a more just and equitable future for our children and grandchildren and for the future for humanity on this, the home of all human beings, our shared Blue Marble.

References

- Bamberger, M., J. Vaessen and E. Raimondo. 2016. *Dealing with complexity in development evaluation: A practical approach*. Newbury Park, CA: Sage.
- Better Evaluation 2020. *Adapting evaluation in the time of COVID-19*. <https://tinyurl.com/2a82ssbj>
- Bitar, K. 2020. If not now, when? Evaluators as agents of change in the time of a global pandemic. Blog post. 1 May. <http://khalil-bitar.com/index.php/blog-4/>
- Chaplowe, S. 2020. Evaluation in the face of the climate emergency: Observations and synthesis. Blue Marble Evaluation blog. 4 May. <https://tinyurl.com/nuu8wh3f>
- Chelsky, J., and Kelly, L. 2020. Bowling in the dark: Monitoring and evaluation during COVID-19 (coronavirus). Independent Evaluation Group blog. 1 April. <https://ieg.worldbankgroup.org/blog/mande-covid19>
- Coates, Ta-Nehisi. 2015. *Between the world and me*. New York: Spiegel & Grau.
- Cowden, John. 2010. Some haphazard aphorisms for epidemiology and life. *Emerging Infectious Diseases* 16 (1): 174–177.
- Duhigg, C. 2020. Seattle’s leaders let scientists take the lead. New York’s did not. *The New Yorker* 4 May: 16–22.

- Euber, J. 2020. Unanticipated crisis lessons: Nonprofit's pivot to online increases participation. NPQ blog. 1 May. <https://tinyurl.com/1odur91g>
- Feinstein, O. 2020a. Dynamic evaluation. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R.D. van den Berg, C. Magro and S. Salinas Mulder, 17–31. Exeter, UK: IDEAS.
- Feinstein, O. 2020b. Dynamic evaluation. Evaluation for Development blog. 28 April. <https://tinyurl.com/8czk7ec7>
- Fortune 2020. The coronavirus economy: How the pandemic is reshaping every aspect of business. *Fortune* May.
- Guterres, A. 2019. UN Secretary-General remarks at opening ceremony of UN Climate Change Conference COP25.
- Guterres, A. 2020. Collaborative leadership is critical to fight both coronavirus and climate change. *Time* 28 April.
- Hayes, D.J. 2020. *Flatten the (climate) curve*. NYU Energy & Environmental Impact Center. <https://tinyurl.com/12xenpcl>
- Heider, C. 2017. *Rethinking evaluation*. Washington, DC: World Bank Independent Evaluation Group. <https://tinyurl.com/rzjrcxfi>
- Hepburn, Cameron, Brian O'Callaghan, Nicholas Stern et al. 2020. Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change? *Oxford Review of Economic Policy* 36 (S1). <https://tinyurl.com/48f9d7k3>
- IDEAS (International Development Evaluation Association). 2019. Prague Declaration on Evaluation for Transformational Change. <https://tinyurl.com/bcm294k7>
- IEG (Independent Evaluation Group). 2020. *Evaluative resources and lessons to inform the COVID-19 response*. Washington, DC: IEG. <https://tinyurl.com/zt5pyixh>
- Karalisi, Thanassis. 2020. Planning and evaluation during educational disruption: Lessons learned from COVID-19 pandemic for treatment of emergencies in education. *European Journal of Education Studies* 7 (4). <https://tinyurl.com/1efk85e8>
- Kolbert, E. 2020. Don't wait. *The New Yorker* 13 January: 13–14.
- Mukherjee, S. 2020. After the storm: The pandemic has revealed dire flaws in American medicine. *The New Yorker* 4 May: 24–31.
- OECD DAC (Organisation for Economic Co-operation and Development Development Assistance Committee) Network on Development Evaluation. 2019. Evaluation criteria. Paris, OECD. <https://tinyurl.com/4x8aszxk>
- Ofir, Zenda. 2020. COVID-19, Part 4. Accelerating. Evaluation for Transformation blog. 20 April. <https://tinyurl.com/3dpcrrdu>
- Patton, M. Q. 2018. *Principles-focused evaluation: The guide*. Los Angeles: Sage.
- Patton, M. Q. 2019a. Transformation to global sustainability: Implications for evaluation and evaluators. *New Directions for Evaluation* 162 (Summer): 103–118.

- Patton, M. Q. 2019b. Climate emergency from a Blue Marble Evaluation perspective. Blue Marble Evaluation blog. December 2. <https://tinyurl.com/37y7qn2q>
- Patton, M. Q. 2019c. Expanding futuring foresight through evaluative thinking. *World Futures Review* 11 (4, December): 296–307.
- Patton, M. Q. 2020a. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Patton, M. Q. 2020b. Evaluation implications of the coronavirus global health pandemic emergency. Blue Marble Evaluation blog. 23 March. <https://tinyurl.com/2bt5bdsh>
- Patton, M. Q. 2020c. Evaluation criteria for evaluating transformation: implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation* 42 (1), 53–89.
- Patton, M. Q. In press. Evaluation ethics in desperate times. In: *Ethics for evaluation* (eds.) Rob D. van den Berg, Penny Hawkins and Nicoletta Stame. New York: Routledge.
- Ramalingam, B., N. S. Singh, A. Mahieu et al. 2020. *Responding to COVID-19. Guidance for humanitarian agencies*. London: ALNAP.
- Rugg, D. 2015. UN spotlights key role of evaluation in new development agenda. *UN News* 13 March. <https://tinyurl.com/4qyd7wtu>
- Rugg, D. 2016. The role of evaluation at the UN and in the new Sustainable Development Goals: Towards the future we want. *Global Policy* 7 (3): 426–430. <https://doi.org/10.1111/1758-5899.12346>
- TEEB (The Economics of Ecosystems and Biodiversity). 2018. *Scientific and economic foundations report*. Geneva: UN Environment. <https://tinyurl.com/1k5icdm0>
- Tolley, L. M. 2020. Learning, evaluating, and staying strong together. AEA365 blog. 3 May. <https://tinyurl.com/4mwuvyty>
- UNICEF (United Nations International Children's Emergency Fund). 2015. *Unless we act now: The impact of climate change on children*. New York: UNICEF. <https://tinyurl.com/4uzdtsfq>
- UNEP (United Nations Environment Programme). 2019. *The United Nations Emissions gap report*. <https://tinyurl.com/1u9nu9oo>
- Vidueira, P. 2020. Aftermath of the 50th anniversary of the Blue Marble (Earth) Day. Blue Marble Evaluation blog. 23 April. <https://tinyurl.com/141qm4wi>
- Williams, B. 2018. DAC evaluation criteria. Evaluation for Transformation blog comments. 7 January. <https://tinyurl.com/1td66m7h>



CHAPTER 3

Evaluation in an Uncertain World: A View from the Global South

ADELINE SIBANDA AND ZENDA OFIR

Abstract. The wicked and intersecting challenges facing the world, brought about by the multiple shocks of COVID-19 and the accelerating impacts of climate change and other effects of the Anthropocene, require that the global evaluation community think and work in fundamentally new ways. It is time that the potential of the Global South – in which we include minority Indigenous societies around the world – is realized, not only because it is most vulnerable to the irresponsible behaviour of societies worldwide, but also because it has important strengths that can help chart a new path for the transformative and sustainable development the world urgently needs. Evaluation will be key in supporting the drastic, much-needed changes. Although exploitation in many different forms persists around the world, the Global South must address the consequences of centuries of colonization, as well as ongoing imperialism and engineered power asymmetries in global governance, economic and financial systems that continue to favour the economically rich Global North. It has become essential that all those who shape and work in the global evaluation system collectively ask how decolonization of the mind and of practice in both the Global South and Global North can be achieved. This is necessary to respect and appropriately attend to the conceptual and methodological experiences, knowledge and wisdom that are deeply embedded in the many diverse cultures of the Global South and that can help advance evaluation in support of systems change and transformation.

An Uncertain World

The year 2020 will always be known as the year in which the world was transformed. The sudden shock of the COVID-19 pandemic brought upon humanity an intertwined health, social and economic crisis of extraordinary proportions that has devastated the physical and mental well-being, education, employment and income of large swathes of populations around the world (UN DESA 2020). It has been a major disruptor of normal life and of business as usual.

It is also an accelerant. What might have taken years has been achieved in months. This is most visible in the movement to digitally distributed work, online retail, education and other services, as well as in the loss of privacy through much more rapid data accumulation by those with the platforms to do so. And although, on the sociopolitical front, the efforts of the Black Lives Matter movement to eradicate racism have made significant gains, so too have the wealth and opportunity disparities and the ideological polarization within and between societies. Reminiscent of feudal and colonial times, in large parts of the world, vast power is now once again concentrated in the hands of a few.

Before the pandemic struck, the 2020 Global Risk Report of the World Economic Forum (WEF 2020a) highlighted the series of complex, interconnected risks facing humanity. The most severe were considered to be socioeconomic, resulting from economic confrontations and political polarization, whereas environmental risks were considered the most serious in the long term, creating a planetary emergency. Other top risks were seen as geopolitical, with tensions resulting from a shift in the balance of power from West to East, and technological, with digital fragmentation and potential information infrastructure breakdown leading to unequal access to the Internet, insufficient global governance and cyber insecurity.

A subsequent study of the short-term effects of COVID-19 indicates an increase in geopolitical concerns due to tighter restrictions on the movement of people, exploitation of the crisis for geopolitical advantage and a reduction in North-South development assistance (WEF 2020b). Technological risk will remain high, with more sophisticated cyberattacks and unforeseen effects of automation. Crucially, the pandemic has had a twofold effect. On the one hand, it has highlighted the urgent need to address the effects of the Anthropocene era, one of which is pandemics such as COVID-19. On the other hand, it has turned attention away from environmental concerns and towards intensive efforts to contain the economic and social fallout of COVID-19. At the time of writing, it is still

uncertain whether the notions of green and circular economies that bring these two dynamics together will be taken up at the necessary scale.

Yet the world remains far from achieving the intent set out in the Paris Agreement. Climate change is occurring more quickly and with greater consequences than expected, while the accelerating pace of biodiversity loss may soon lead to the disruption of entire supply chains and the collapse of food and health care systems around the world (UN DESA 2020). The interconnected nature of the global risks, the magnitude of the 'black swan' pandemic and accelerating challenges worldwide mean that strategists, decision makers and ordinary people going about their daily business face unparalleled uncertainty under rapidly changing circumstances. If the world is to address the major challenges that confront humanity in this century, the capacity to address unpredictability will be an essential part of the quest for greater resilience in mindsets, behaviours and the institutional systems that hold society together.

An Entangled World

The COVID-19 pandemic has graphically demonstrated the connections between actions and events around the world – and thus the interdependence between people and the social-ecological systems on which humanity depends. This has several major implications for efforts to solve the new and intractable challenges facing humanity today.

We need a systems view of the world to inform change strategies. 'Like a double helix, the SDGs [Sustainable Development Goals] and the COVID-19 pandemic responses are intertwined and cannot be tackled by a piecemeal approach' is how the United Nations Development Programme recently articulated the need for a complex systems-informed approach to responses and solutions (UNDP 2020). Reinforcing the ambitious agenda underlying the SDGs, integrated solutions are considered crucial for building a greener, more inclusive future. Actions and challenges affect one another across country borders and from a global to a local level; a Blue Marble perspective on the world's societies and ecosystems (Patton 2019) is imperative, and this must be reflected in practical responses and solutions.

We need concerted collective action. Sustainable development has become the responsibility of all countries in the world, whether economically rich or poor. At the same time, geopolitical strife and political myopia have the potential to lead to rapid deglobalization and a shift away from cooperation across geopolitical and ideological boundaries. The crises that

define the era of the Anthropocene demand a revisiting of concepts and values that undergird development. 'We cannot go back to the way it was and simply recreate the systems that have aggravated the crisis. We need to build back better with more sustainable, inclusive, gender-equal societies and economies' (Guterres 2020). Efforts to recover from the devastation of the pandemic are likely to weaken commitment to the most vulnerable and might delay action towards a greener, more equitable future. Priorities will have to change, and clear articulation is needed of the values that will help determine the appropriate balance between people and the planet, as well as between the economic, social, environmental and technological aspects of progress articulated in the 2030 Agenda and Paris Agreement.

The Divide: Global South, Global North

[T]he story we've been told about rich countries and poor countries isn't exactly true. In fact, the narrative we're familiar with is almost the exact opposite of reality. There is a very different story out there, if we are willing to listen to it. –Jason Hickel, The Divide (2017)

At present, all indications are that fallout from the pandemic will severely affect the majority of the world's population, especially the more than 6 billion people in some 140 countries commonly seen as 'developing' countries, or the Global South (UN DESA 2020). Some argue that it is inappropriate to divide the world into the Global South and Global North¹. There is, they say, too much social, economic and cultural diversity within each, as well as rising incomes in some poor countries, increasing inequalities in many rich countries and increasingly fluid political alliances around the world. Yet disregarding the differences between these two parts of the world has led to many false narratives about development progress, priorities and strategies.

¹ The distinction is not geographic, but socioeconomic and political. The economically rich countries are incidentally nearly all located in the northern hemisphere; the economically poorer ones are mostly, but not exclusively, south of the equator. The 'Global North/Global South' terminology has recently become more popular as a result of the backlash against the notion that some countries are 'developed' and therefore superior, a state that the others should strive towards and that can set standards for others. In turn, the developed/developing country terminology replaced the even more patronizing notion of a First, Second and Third World.

The differences between the Global South and North are in part historical, with solidarity among countries in the South resulting from a shared history of colonization, marginalization and disempowerment. In his book *The Divide*, Jason Hickel (2017) dissects the deliberate creation of poverty and what he calls the ‘economics of planned misery’. Global indexes and analyses continue to highlight the many persistent disparities between these two parts of the world while book upon book and document upon document have revealed how these disparities are largely the result of the odds having been stacked against the South (Chang 2008; Moyo 2009; Stiglitz 2002).

It is likely that the COVID-19 pandemic will further inhibit the efforts of the most fragile countries to escape their enduring poverty traps. According to estimates, it might push half a billion people around the world back into poverty (Oxfam 2020) and 71 million into extreme poverty (UN DESA 2020); some 1.6 billion people in the informal sector are facing destruction of their livelihoods (ILO 2020). This will put exceptional strain on many countries – most of which will be in the Global South given their often relatively weak institutions, limited access to technology, frequent political instability and greater vulnerability to human-induced disasters.

The Urgency of Transformational Change

The UN and member states are sleepwalking towards failure...it is time to acknowledge that the SDGs are simply not going to be met. –Philip Alston, The Parlous State of Poverty Eradication (2020)

The magnitude and nature of the risks facing humanity in 2020 leave little doubt that, without extraordinary action, the chance of achieving the SDGs by 2030 is fading quickly (UN DESA 2020). Transformational change – drastic, large-scale change that fundamentally changes the structure of a system so that it will not return to its prior state – has been a central part of the 2030 Agenda. It is seen as essential to achieving the targets of the Paris Agreement. No less important for some, transformational change is also seen as a change in systems to support social justice and inclusion within and between societies (Mertens 2008).

Both the Global South and North carry responsibility for transformational change but from two very different vantage points. The economically rich countries in the Global North must curb the planetary boundary overshoot that has resulted from their patterns of overconsumption (Raworth 2017; UN DESA 2020), curb the injustices and inequalities in their own

countries, change strategies that harm other societies and contribute to equal opportunities for all. In the Global South, governments must continue to work towards escaping poverty traps but are now also under unparalleled pressure to deliver services, provide social protection and ensure social cohesion under highly challenging circumstances (UN DESA 2020). Adaptability, agility and resilience are essential for appropriate responses to these challenges and risks. This means that many countries in the Global South with less educated populations, more fragile institutions, fewer financial resources and less power are in a weaker position to effect urgently needed positive transformations.

What can and should evaluation contribute in this situation?

Evaluation for This Time

The era of the Anthropocene demands that evaluation follow a new evolutionary path. It must accompany and assess responses to ongoing transformations such as those resulting from the COVID-19 pandemic. It must also support interventions aimed at triggering and shaping urgently needed transformations, such as mitigating the effects of climate change. To do so effectively, evaluation itself, and the system that shapes it, needs to transform. A free-standing practice that has been evolving slowly from project and programme evaluation in the West – one characterized and shaped by the preferences of clients, the pressure to show quick results and simplistic use of performance measures and indicators (Picciotto 2020) – will not be suited to the demands of this era.

What Thomas Schwandt (2009) calls the Western evaluation imaginary – the way in which we collectively see evaluation in the contemporary western world – provides for the common understanding and hence practices that have given evaluation an Enlightenment bias and made it ‘thoroughly modernist’. Although the globalization of evaluation over the past two decades has opened it up to many different influences, development evaluation in particular remains in essence a Western practice (Chouinard and Hopson 2016; Waapalaneexkweew/Bowman-Farrell 2018) in spite of early caution about the belief in Euro-Western universalism and practices in evaluation, increasing awareness of the importance of culturally responsive evaluation and exemplary contributions by Indigenous and Black evaluators (see e.g. Chilisa 2017; Cram et al. 2015; Cram, Tibbetts and LaFrance 2018; Hood, Hopson and Frierson 2015).

Brouselle and McDavid (2020) echo the notion that evaluation is not ready for the Anthropocene and point out that ‘most evaluators think in a

micro context, a legacy of evaluation practice that serves other disciplines, decision-makers, policy-makers, funding agencies and beneficiaries'. The few evaluators working in climate adaptation and mitigation (see e.g. Rowe 2019; Uitto, Puri and Van den Berg 2017) and the newly developed Blue Marble Evaluation approach (Patton 2019) have contributed most, and most visibly, to Anthropocene-oriented evaluation practice – focusing on sustainable development as large systems change that connects people and nature.

During a keynote speech at the 13th Biennial Conference of the European Evaluation Society, Schwandt (2019) noted the emergence of post-normal evaluation, which is based on new ways of thinking and forms of evaluation practice that reflect 'assumptions of unpredictability as well as incompleteness, instability and a plurality of perspectives in value determination'.

All of this points at least in part to new frontiers for evaluation. Realist Evaluation, Developmental Evaluation, Dynamic Evaluation, Principles-Focused Evaluation and Blue Marble Evaluation already provide a sound basis for further work from this perspective, but the political economy of evaluation; focus on bureaucratic and indicator-driven, results-based management and lack of professionalization and capacity deficits that constrain evaluation (Picciotto 2020) also work against uptake of these approaches in the evaluation system beyond a focus by some on experimentation, adaptive management and adaptive learning. Complexity-oriented frameworks and systems-informed methodologies are emerging but have as yet only limited application in practice.

Evaluation in service of transformational change can, on the one hand, relate to evaluation as a critical voice in environments marked by significant inequalities and power differentials, in line with Mertens's (2008) notion of transformative evaluation and, on the other, in environments marked by uncertainty, ambiguity and interpretability. In both cases, this places evaluation in the role of wayfinding (Schwandt and Gates 2016). It is necessary to push evaluation professionals not only to answer the question 'What makes this the right thing to do?' but also to engage with 'What makes this the right thing to do if evaluation is to contribute to its full potential to the challenges the world faces?'

Evaluation scholarship that can enhance theory and effectively translate it into practice is urgently needed to advance post-normal evaluation in service of transformation. More than that, evaluation professionals in the Global South need to determine what the systematic practice of evaluation would have looked like in terms of concepts, theories and practices if it had not been invented and advanced primarily in the Global North – and with a lens that can best support efforts to inform transformational change at

this time. Although this is unlikely to change key aspects of the valid critique against current practices that beg for professionalization (Picciotto 2020), it will enhance our understanding of how far we can stretch the contours, boundaries, limitations and value proposition of the field in different parts of the world. More than that, we believe that the potential for supporting transformation through evaluation, and for transforming evaluation through a focus on transformation, can benefit from emphasis on the innovations that lie silently in world views, philosophies and traditions in the Global South.

Power Asymmetries in the Global Evaluation System

Long-standing patterns of power that have emerged as a result of colonialism have influenced the evaluation field. Haugen and Chouinard (2019) note that power is complex, intangible and invisible. The purpose of evaluation is often to help donors or development agencies justify their existence. Funders and commissioners have the power to make decisions about who participates in evaluations, what is evaluated and what data are collected – often marginalizing communities that may have different priorities and no means to implement the development projects they want for themselves (Chilisa 2015). The power dynamics continue between international and national evaluators; national evaluators are often downgraded to research assistants, regardless of their qualifications. These power asymmetries often go unchallenged. Evaluators from the Global South often unquestioningly accept terms of reference because of their own colonized mindsets, which leads them not to question the value or merit of their assignment, regardless of their high level of experience, or they are reluctantly forced to accept the situation because they need the money or want to develop their profile and capacity.

Colonization and Decolonization

The power asymmetries in the evaluation system can best be understood through the experiences and legacies of colonialism in the Global South.

Stages of Colonization

Hoppers and Richards (2011) refer to phases of colonization. The first phase was the colonization of lands and physical spaces, followed by the colonization of the mind through education and other societal systems.

Achieving independence from the colonial administration did not remove all stages of colonization; colonization of the mind through education and the way development and development evaluation are practiced still remain. Maldonado-Torres (2007) notes that persistent patterns of power that emerged as a result of colonialism continue to define culture, labour, intersubjective relations and knowledge production well beyond the colonial administrations. Colonialism is maintained alive in books, criteria for academic performance, cultural patterns, common sense, self-image of peoples, aspirations of self and so many other aspects of our modern experience.

Education systems still based on Western curricula and methods of education have influenced views of development and evaluation held in the Global South (Chilisa 2012; Gaotljobogwe et al. 2018). This is further entrenched in the minds of Global South evaluators. What are currently promoted as ideal development models are measured using yardsticks from the Global North, and anything conceived and designed by the Global South is said not to meet the 'standards' (Chilisa 2015).

About Decolonization

Decolonization is the change that colonized countries go through when they become politically independent from their former colonizers, but decolonization is not merely a matter of political independence. Structures of government and other institutions, how post-colonial countries are economically organized and how people in those countries are encouraged to think are often still determined by the former colonial powers because of the economic and cultural power that they wield. To overcome this legacy, it is essential to decolonize the intellectual landscape of the country and, ultimately, to decolonize the mind of the formerly colonized (Oelofsen 2015; Gaotljobogwe et al. 2018).

Such is the case with evaluation. If it is to be truly transformative, the values, power dynamics and intellectual landscape that have shaped the evaluation field today must be questioned. Going below the surface and beyond the symptoms to address the deep causes for the state of evaluation, we must look closely at the layers of complexities within the silenced communities that are often the subject of our inquiry.

Decolonizing Systems

To do this, we must think in systems. The process of decolonization is not a one-time event. It is a process that can take decades, across multiple

sectors and spheres of life, within and outside the international systems that direct and shape evaluation. Such a process is also resisted not only by the colonizers, but also by those from the Global South whose minds are still bound by the chains of colonialism (Lent 2017; Ikuenobe 2017).

Within the evaluation context, this requires a twofold focus. Decolonization means confronting and challenging the colonizing practices that have influenced evaluation in the past and are still present today. It implies questioning subtle colonial expressions in evaluation and development practice, for example definitions of 'households' or 'beneficiaries', forced participation, indicators of empowerment, ethical procedures and the paradigms currently shaping the practice. Chilisa (2015, 14) argues that

Decolonisation of evaluation can be viewed as the restructuring of power relations in the global construction of evaluation knowledge production, such that the African people can actively participate in the construction of what is evaluated, when it is evaluated, by whom and with what methodologies.

The way evaluators have aligned themselves with world views, theories and practices that have emanated from specialists in the Global North has diluted their impact and potential for transformation by reinforcing the existing way in how 'development' is viewed and 'done'. In our experience, there are zealous gatekeepers on both sides, policing and suppressing any move away from the current state of affairs. These same gatekeepers often express their resistance by questioning, for example, whether it is possible to ensure rigour and reliability in the data collection and analysis processes if Indigenous methodologies are used. This argument reveals the importance of the world views or paradigms used in characterizing 'rigour and reliability'; they are often expressed from a post-positivist paradigm rather than from an Indigenous, relational paradigm.

Ndlovu-Gatsheni (2013) argues that decolonization alone will not work. We need a simultaneous process of decolonization *and* de-imperialization – letting go of the egotism that reproduces and maintains the idea that the Global North has everything to teach the rest of the world and nothing to learn from other people and their civilizations. These processes require effective dialogue between the colonizers and the colonized. In addition, de-imperialization is necessary to dismantle the racially hierarchical modern world system and reconstruct the asymmetrical power relations in this system. Decolonization must be deepened to address deep cultural, psychological and epistemological issues.

The Power to Transform Our Future Comes from Within: Examples from the Global South

The responsibility to address the power asymmetries in evaluation and to embrace and promote evaluation scholarship from the Global South lies with both Global North and Global South evaluators. The power to transform is within all of us. We therefore must transform from within to transform the world that we live in. It is about shifting our mindsets – decolonizing and de-imperializing the mind.

Philosophies in East and Southeast Asia

For millennia, philosophies such as Buddhism, Taoism and Confucianism have shaped the world views, traditions and actions of populations in Southeast and East Asia. Buddhism, an ancient philosophy developed around the 6th century BCE, lends itself to interpretations of common concepts and practices in evaluation that differ significantly from the norm (Dinh, Worth and Haire 2019; Russon 2008; Russon and Russon 2010). Among others, the trigrams in the highly influential *I Ching*, or *Book of Changes*, reflect combinations of yin and yang – concepts related to the interdependence and complementarity of nature and mind, of the cosmos and humans, of natural events and political systems; it is perceived that opposing the natural order of the universe will lead to disaster and destruction.

Furthermore, Buddhist understanding of causation is that all phenomena exist as a result of the interaction of multiple factors. Cause and effect cannot be isolated; they arise together and are recursive. This interdependence means that the observer of a process of cause and effect cannot be isolated from the process itself (Russon 2014), that there is a middle way between determinism and uncertainty and that intrinsic validation rather than empirical verification is required for evidence (Dinh, Worth and Haire 2019). Buddhism also perceives the world as being in a state of transient flux, with endless processes of change, which means that ‘impact’ as we define it does not exist – only ‘a combination of conditions that come together in a certain way at a certain point in time’ (Russon 2014).

Dinh, Worth and Haire (2019) applied Buddhist principles to the most significant change technique and found them compatible. Evaluation rooted in Buddhism will be well aligned with systems thinking and complexity science, as well as with the essential values that are being advanced to address the challenges of the Anthropocene. There are also clear parallels between these principles and the world views and approaches of other

Indigenous societies in the Global South and around the world. Buddhist values and principles also underpin the Gross National Happiness Index and philosophy that have steered Bhutan's policies over past decades towards making it the world's only carbon-negative country (Yanka 2018).

China's Transformative Development

China has transformed in just four decades. It has been responsible for lifting some 850 million people out of poverty – nearly all the progress in reducing poverty worldwide during that period (Alston 2020; World Bank 2018), yet it did not follow any dominant narrative about development such as the neoliberal structural adjustment programmes that forced economically poor countries in the 1980s and 1990s to reverse their impressive economic gains after the end of colonization (Chang 2008; Hickel 2017; Moyo 2009; Terreblanche 2014). Detailed analyses (Ang 2016) have shown that some of the key reasons for China's success have been its treatment of development as a complex adaptive system, anchoring its five-year development plans in concepts such as directed improvisation and co-evolution and using evaluative practice in defined ways. It has been explicit about the values and principles that guide its policies and programming. It has also illustrated graphically that development narratives and models based on Euro-Western conventional, often reductionist wisdom can be inappropriate and even destructive, yet we apply them unquestioningly in theories of change and in evaluation frameworks.

Indigenous Societies Worldwide

Over the past two decades, Indigenous evaluators from minority populations in New Zealand, Australia, the Pacific and North America² have made great strides in bringing alternative framings for evaluation theory and practice to the fore in support of culturally responsive and culturally responsive Indigenous evaluation. The underlying philosophies apparent in these frameworks reflect a complex adaptive systems view of the world, with values shaped by social justice and a holistic perspective that honours relationships between people, as well as between humanity and nature, in

² Although many of these countries are generally considered to be part of the Global North, their indigenous communities have had experiences similar to those in the Global South.

line with current notions of sustainable development – all essential for solutions to the challenges of the Anthropocene.

Cram, Pipi and Paipa (2018) note that Indigenous ways of knowing and evaluation are ‘holistic, relational, and oriented to a place of dwelling’; build trust based on respectful practices and reciprocity and honour emergent, self-determining processes framed by Indigenous evaluators and owned by respected communities. Tyson Yunkaporta (2020) demonstrates how well aligned Indigenous wisdom in Australia is with current understanding of global systems and complexity concepts – emphasizing relationships between people and with nature, the importance of living within the pattern of creation and ensuring perspectives different from conventional dominant frameworks and narratives to resolve contemporary challenges.

You would think a complex system like a marketplace would be able to interact more dynamically with complex ecological systems. This kind of dialogue always breaks down, however, when it is mediated by the cult of reductionism... Perhaps a first step would be a subtle shift in the focus of inquiry to include an Indigenous orientation, examining multiple interrelated variables situated in place and time (Yunkaporta 2020, 170).

Indigenous framings of life and hence of evaluation tend to be circular and cyclical. Life is a state of constant flux, and change is part of the natural cycle of life, a ‘two-way process of interaction with the universe’ (Mustonen and Feodoroff 2018, 110). The Athabascan Circular Model (or Seasonal-Cyclical Model) that Alaska Native peoples use reflects this logic. Anderson et al. (2012, 577) describe how Debbie, an Indigenous workshop participant, articulates her sense of theory of change when shifting from a linear to a cyclical model:

A circle is a model just about any Indigenous person can relate to. Linear thinking does not make sense to me. How can something end and not begin something else?...many non-native people complain no end about winter coming; the cold weather, the darkness. Without winter, how would we have spring and summer? I do not understand the complaining but it happens, and it irritates me.

A multidimensional, culturally responsive Indigenous evaluation model based on Stockbridge-Munsee/Lunaape traditional teachings acknowledges the importance of situating living (and hence evaluation) within a broader historical context and the entangled effects and value of the lived experience of interaction – including colonization. It encourages the resolve to enable and restore balance through contributing, developing

relationships and sharing with others and to honour new beginnings and individual and societal development based on experience, knowledge and wisdom from within and from others (Waapalaneexkweew/Bowman-Farrell 2018; Waapalaneexkweew/Bowman-Farrell and Dodge-Francis 2018).

The world of the Māori in Aotearoa (New Zealand) is also about 'kinship with other people, with our environment, and within the cosmos' (Cram et al. 2015, 306). Kaupapa Māori evaluation legitimizes and recentres Māori reality and view of the world and asserts their right to conduct evaluations that are by, with and for Māori; evidence generated in a Māori context by outsiders will, therefore, not be considered credible unless permitted in a supportive role.

Ubuntu in Africa

Global South communities are often bound by traditional belief systems, values and cultural practices passed from one generation to another. In the case of Sub-Saharan Africa, this is embraced in the concept of *ubuntu*, a Zulu term that means 'I am what I am because of who we all are'. Ubuntu is based on a set of explicit values that includes caring, sharing, reciprocity, cooperation, compassion and empathy – recognizing that, to develop, flourish and reach their full potential, human beings need to conduct their relationships in a manner that promotes the well-being of others and the environment (Mawere and Van Stam 2016). The values championed in *ubuntu* have informed and shaped African cultural, social, political and ethical thought and action (Mawere and Van Stam 2016). Studying languages that include the idea of *ubuntu* shows that it is more than just a word; it is a lived reality. It is what defines 'being African'³. Ubuntu values emphasize well-being and the environment.

The notion of *ubuntu*, oriented towards communal or societal rather than individual interests and grounded in a sense of relationships that include nature, reflects similar concepts in other societies in the Global South and in Indigenous communities in the Global North. For example, the Aztecs, whose descendants still live in Latin America, practised socially centred virtue ethics centred on the cooperative, on family and friends and on the collective rituals and routines of daily life (Purcell 2018). Confucianism in China (revived as New Confucianism in the 20th century) has as fundamental principles the notions of *ren* or 'humaneness', which emphasizes harmony

³ Khomba (2011) explores a number of African languages with *ubuntu* derivatives, showing the breadth of *ubuntu*.

in society, and *li*, which embodies the web of interactions between humanity and nature. In Māori society, concepts such as *whānau* and *whanaungatanga* emphasize the importance of extended family or kinship, societal collectives and establishing and maintaining relationships in society (Cram et al. 2015; Paipa et al. 2015).

South–South Cooperation

A good example of evaluative approaches reflecting the Global South/Global North divide can be found in efforts to develop monitoring and evaluation frameworks that recognize the different values and principles underpinning South–South and North–South international cooperation (Besharati, Rawhani and Rios 2017). Both include a focus on results, but whereas North–South development assistance tends to focus on process qualities such as local ownership, mutual accountability, harmonization, alignment and inclusive partnerships⁴, South–South cooperation principles have a very different tone – valuing solidarity, partnerships between equals, national sovereignty and ownership, non-interference in domestic affairs, mutual benefit or win-win and fostering collective self-reliance (UN 2009). Efforts to find common ground at the high-level meetings of the Global Partnership for Effective Development Cooperation became mired in controversy (IDS 2018) because many countries in the Global South do not want to let the criteria, measures and methods of North–South development assistance dictate how South–South cooperation is understood and its performance measured.

Global Risks from a Global South Perspective

A review of the global risks referenced earlier in the chapter reveals that some risks and their impacts might result from the fact that the relevance for, and ways of knowing of, people from the Global South are not taken into consideration; this represents a certain form of colonization. We discuss here two of these risks, using the concept of ubuntu, to illustrate the power dynamics at work and the neglect of Indigenous knowledge and social norms, culture and historical perspectives.

⁴ As determined at the Paris Declaration on Development Effectiveness in Paris in 2005, the Accra Agenda for Action in 2008 and the Busan Outcome Document in 2011.

Technological Risk

When we consider data and technology, inequalities and power asymmetries are more powerfully reflected in the Global South than the Global North. Data are often assumed to be accurate and neutral, good for evidence in the evaluation profession, but this is not always the case. Sensitive evaluation data about vulnerable populations can be stolen or misused in cyberattacks, or sensitive on-line meetings can be disrupted, especially now that most discussions and interviews have moved on-line. Data and technology can be used in ways that harm people at critical moments; for example, psychometric tests may lead to discrimination against certain groups of people during major life events such as going to university, borrowing money or getting a job or promotion. These data systems create a vicious circle, feeding on each other (O'Neil 2016). Poor people are more likely to have bad credit and live in high-crime neighbourhoods, surrounded by other poor people. The result is that poverty is criminalized.

Thus, contrary to the belief that tools and technologies are scientific and fair, algorithms can be biased or racist, reflecting the make-up and beliefs of the designer. Major software companies in the North also tend to develop and test algorithms based on averages of data taken from mainly white men (O'Neil 2016). A completely different picture will be produced if the diversity in society is taken into account. At a MerITech conference in 2019, it was noted that some facial recognition systems misidentified black people 5 to 10 times as often as white people.

It is a tragedy that many continue to use such data as the Truth. These systems define their own reality and use it to justify their results. When evaluators use these data in assessments, the evidence will be inaccurate. This situation is of particular importance in the Global South, which largely depends on software from companies in other parts of the world. In the aftermath of the COVID-19 pandemic, the use of remote sensing, (automated) online data collection and big data will accelerate, which will exacerbate data interpretation challenges in the Global South, where accurate data – and the nuanced data derived from qualitative information – are needed.

This is contrary to the traditional values undergirding much of the Global South; for example, ubuntu emphasizes caring, sharing, reciprocity, cooperation, compassion and empathy. Technology and data experts very often do not involve Global South societies, including Indigenous populations, in design, testing and feedback mechanisms. There is no co-creation, cooperation or caring about how the data and technology can prejudice their whole being and affect them for the rest of their lives.

Environmental Risk

From the concept and values of ubuntu, we note the emphasis on well-being and environment. Africa had a history of conservation before colonization. In southern Africa, people have traditionally lived in harmony with nature, and the philosophies of sustainability and conservation were inherent in their societies. As hunters and gatherers, Indigenous people depended on natural resources for survival.

Every human society has its own unique culture that is closely related to the environment and nearby natural resources; culture and biodiversity conservation are tightly interwoven because, as the environment changes along with the availability of resources, the culture of the affected people may also change. Through cultural practices, local people acquired valuable knowledge and skills for the conservation of biodiversity and sustainable use of natural resources, yet the conservation practices of the past (and present) often cast doubt upon traditional systems of biodiversity conservation and labelled them as unscientific and unreliable (Ministry of Environmental Affairs 2008). It is important to question these traditional systems of biodiversity conservation to understand how they managed, monitored and evaluated their environment to ensure sustainable use of resources.

Before colonization, there were strategies in the Global South to address environmental risks and impacts. In Africa, tradition, culture and religion influenced resource use, and there were controls to reduce depletion of resources. Land use was managed through controlled access to natural resources. Among other influences, resource use was influenced by religious beliefs and affected by local traditional healers; there were traditional prohibitions against killing certain species such as hyena or chameleon; people were prohibited from hunting or eating their totem animals, such as lions, eland, zebra, monkeys, squirrels and crocodiles (Ministry of Environmental Affairs 2008) and areas were demarcated for specific purposes, possibly to prevent overuse of these species. Some areas were set aside for religious purposes, including sacred forests, burial sites and hills for ceremonies and rituals. There were also areas that were not inhabited, to prevent degradation of fragile soil (Department of Environment Affairs 2008).

These practices changed during colonization for many southern African countries. In Zimbabwe for example, Indigenous people were moved to unproductive land or reserved areas and sometimes to lands that were fragile and unsuited for human habitation or large populations, yet when current conservation or environmental efforts are evaluated, the history of the communities and their cultural heritage are often not taken

into consideration. This leads to conclusions that communities are causing land degradation because of overpopulation or because they live in watershed areas or fragile soils. There is often little examination of how they came to be there in the first place and often no attempt to understand their side of the story and how Indigenous knowledge or their ways of knowing can be tapped into to find solutions.

One may ask why it is this way. Power relations are a major factor. The person who commissions or funds the evaluation has the power over what is done and how. The commissioners of the evaluation have an important voice, whereas the communities in need of 'development' are often perceived as not knowledgeable enough to craft their own destiny or have an independent, authoritative voice in plans and implementation activities. Global sustainability is rarely addressed using Indigenous world views and perspectives. Discussions of Indigenous knowledge systems tend to be polite acknowledgements of connection to the land rather than true engagement (Yunkaporta 2020).

Evaluation and Social, Political and Economic Inequalities

Evaluation design is a highly challenging task influenced by political factors and implementation realities. Social, political and economic inequalities in access to natural resources and public programmes on health, education and livelihood securities need context-sensitive approaches combined with diversified tools and techniques. Development is not only a question of economic betterment, but is also about improving people's sense of belonging, self-realization and hopes for the future. Similarly, evaluation is not only about measuring goal attainment and impacts, but also about understanding the sociocultural fabric of society and the bridging of gaps in unequal opportunities. This has led to multiple ongoing efforts in Latin America, Asia, Africa and other parts of the world to develop tailored monitoring and evaluation approaches and methods that account for the intersections of gender, culture and rights.

Stimulating Evaluation Practice in the Global South

The South-to-South Evaluation Initiative was developed in 2018, building on two decades of committed efforts to identify and address asymmetries

in evaluation in the Global South. Leaders of five regional voluntary organizations for professional evaluation in the Global South have spearheaded the initiative⁵, with the goal of traversing 'the last mile' in addressing asymmetries in power, decision-making, resources and knowledge in the evaluation ecosystem in the Global South to contribute to the sustainable development of all nations. It has been raising awareness and advocating for early adopters to join the initiative, and some of the engaged regional associations have started to inspire stronger evaluation scholarship and research on evaluation in the South.

One such example is the Made in Africa Evaluation initiative that the African Evaluation Association is implementing. It is an evaluation agenda that is increasingly prioritizing evaluation for transformative development based on evaluation frameworks and techniques – rooted in African world views and Africa's development vision and priorities and aimed at inspiring respect for human dignity yet fully engaged with international evaluation practice. The Made in Africa Evaluation has an implicit theory of change that reflects world views and frameworks that recognize the interconnectedness of people and their environment and is rational yet mystical and spiritual (Chilisa 2015; Gaotljobogwe et al. 2018). It has recently mobilized financing for a set of research initiatives on evaluation from Global South perspectives in an effort to enhance scholarship and study in this field of work.

The Made in Africa Evaluation is intended to challenge the prevalent practice of designing evaluation approaches and tools without attending to cultural responsiveness or to the diversity that manifests itself in the plethora of cultures, religions, languages and histories on the continent and in gender and ethnicity. It is also set to challenge the extractive nature of evaluation practices that fail to benefit those in the society who give their time and expertise for the sake of the evaluations. It must question the perennial success stories told about interventions when realities on the ground are completely different. It is encouraging greater engagement with and recognition of African data collection methods such as storytelling, folklore, music, dance, oral traditions and the use of African languages (Chilisa 2015). The African Evaluation Guidelines have also been revised using a Made in Africa Evaluation lens.

⁵ African Evaluation Association; Asia Pacific Evaluation Association; Caribbean Evaluators International; Community of Evaluators South Asia; and Latin-American and Caribbean Monitoring, Evaluation and Systematization Network

Business Unusual: Mobilizing for the Future

The major shifts the world needs to respond to the challenges of the era of the Anthropocene also require a major shift to 'business unusual' by the evaluation sector. The sudden shock of the COVID-19 pandemic has affected the whole of humanity. It knows no borders and has made everyone vulnerable, yet it has also brought solidarity and galvanized the world, demonstrating that drastic change on a global scale is possible. This willingness to act should encourage the global evaluation community to increase the intensity with which it addresses power asymmetries and injustices in global systems in general and in evaluation ecosystems in particular. We should collectively draw vigorously and respectfully from the diversity of experiences and knowledge systems available. Neglecting to ensure that all parts of the world contribute substantively to evaluation depletes its transformative potential.

What and how we evaluate, who we work with, how well our work empowers people in the Global South and how we respond in real time to global catastrophes – climate change, pandemics, biodiversity loss, overconsumption and wars, all increasingly leading to profound social instability and suffering – must change. We must find appropriate spaces for all world views and knowledge assets – even though few academic centres are focusing on evaluation in the Global South while good evaluators in the field are constantly overworked. Because this inevitably diminishes the chance of creativity and innovation, solutions must be found.

Evaluation professionals and communities and societies in the Global South should be consistently valued – their ways of knowing solicited and respected, their practices embraced and encouraged – whether they are working in the Global South or the Global North. It is essential to question narratives, models, frameworks and theories of change about development and evaluation imported from the Global North. We must study, synthesize and make visible insights and narratives from different knowledge systems about the nature of concepts such as change, causality and evidence. We must rethink the values and approaches that underlie assessments of progress, success and impact, and we must insist on making cultural responsiveness part of all terms of reference. We must conduct, synthesize and learn from evaluations of South–South cooperation and innovative approaches to sustainable development while working together to advance South–South evaluation. We should develop mechanisms for intensive generation, documentation and accumulation of innovative Global South approaches to the challenges the world faces and update education curricula and short-term training with Global South–developed theories and practices.

If the field of evaluation is to help facilitate transformational change on a large scale, evaluation professionals must use and expand on the principles laid out in *Blue Marble* – developmental and principles-focused evaluation – but it must also pay special attention to incorporating complex adaptive systems concepts, power dynamics in systems and explicit values such as those espoused through the notion of, for example, ubuntu (caring, sharing, reciprocity, cooperation, compassion and empathy) – values that will help us respect, collaborate, co-create and empower.

The crises facing humanity offer us challenges but also opportunities to do the best we can in our professional practice. Evaluation specialists worldwide must stop enclosing themselves in a cocoon as if only the evaluation world matters, and they must learn from other disciplines, sectors, fields of work and worlds of financing. New types of partnerships for collective action, South–North as well as South–South, will be essential to create transformation systems in evaluation – that is, change agents aligning to work in synergy towards common goals.

Such actions will demand commitment, a sense of urgency and robust positioning of evaluation in the global sustainable development agenda. They will also require renewal in the global evaluation agenda, centred on the demands of the Anthropocene and the effects of the COVID-19 pandemic. If evaluation professionals – financiers, commissioners, managers, educators, researchers and evaluators – across the Global South and Global North share this responsibility, working in tandem and on equal footing for the benefit of all, we may well become the best that the field of evaluation can offer the world at this critical juncture.

References

- Alston, Philip. 2020. The parlous state of poverty eradication. Report A/HRC/44/40 of the UN Special Rapporteur on Extreme Poverty and Human Rights. Paper presented at the Forty-Fourth Session of the Human Rights Council, 2 July.
- Anderson, Clara, Malinda Chase, James Johnson III et al. 2012. It is only new because it has been missing for so long: Indigenous evaluation capacity building. *American Journal of Evaluation* 33 (4): 566–582.
- Ang, Yuen Y. 2016. *How China escaped the poverty trap*. Cornell Studies in Political Economy series. Ithaca, NY: Cornell University Press.
- Besharati, Neissan, Carmel Rawhani and Ornela G. Rios. 2017. A monitoring and evaluation framework for South–South cooperation. Network of Southern Think-Tanks, Africa Chapter.

- Brousselle, Astrid, and Jim McDavid. 2020. Evaluators in the Anthropocene. *Evaluation* 26 (2): 190–204.
- Chang, Ha-Joon. 2008. *Bad Samaritans: The guilty secrets of rich nations and the threat to global prosperity*. New York: Random House.
- Chilisa, Bagele. 2012. *Indigenous research methodologies*. Los Angeles: SAGE Publications.
- Chilisa, Bagele. 2015. *A synthesis paper on the Made in Africa Evaluation concept*. Accra: African Evaluation Association.
- Chilisa, Bagele. 2017. Decolonising transdisciplinary research approaches: An African perspective for enhancing knowledge integration in sustainability science. *Sustainability Science* 12: 813–827. <https://doi.org/10.1007/s11625-017-0461-1>
- Chouinard, Jill A., and Rodney Hopson. 2016. A critical exploration of culture in international development evaluation. *Canadian Journal of Program Evaluation* 30 (3): 248–276.
- Cram, Fiona, Vivienne Kennedy, Kirimatao Paipa et al. 2015. Being culturally responsive through Kaupapa Maori evaluation. In: *Continuing the journey to reposition culture and cultural context in evaluation theory and practice*, (eds.) Stafford Hood, Rodney Hopson and Henry Frierson, 289–311. Charlotte, NC: Information Age Publishing (Evaluation and Society series).
- Cram, Fiona, Kataraina Pipi and Kirimatao Paipa. 2018. Kaupapa Maori evaluation in Aotearoa New Zealand. *New Directions for Evaluation* 2018 (159): 63–77.
- Cram, Fiona, Katherine A. Tibbetts and Joan LaFrance. 2018. Editors' notes: A stepping stone for Indigenous evaluation. *New Directions for Evaluation* 2018 (159): 7–16.
- Dinh, Kathryn, Heather Worth and Bridget Haire. 2019. Buddhist evaluation: Applying a Buddhist world view to the most significant change technique. *Evaluation* 25 (4): 477–495.
- Gaotljobogwe, Michael, Thenjiwe E. Major, Setlthomo Koloi-Keaitse et al. 2018. Conceptualising evaluation in African contexts. *New Directions for Evaluation* 2018 (159): 47–62.
- Guterres, António. 2020. Opening remarks to the press at the launch of the UN Comprehensive Response to COVID-19, 25 June. www.un.org/sg
- Haugen J.S., and Jill A. Chouinard. 2019. Transparent, translucent, opaque: Exploring the dimensions of power in culturally responsive evaluation contexts. *American Evaluation Journal* 40 (3): 376–394.
- Hickel, Jason. 2017. *The divide: A brief guide to global inequality and its solutions*. London: Penguin Random House.
- Hood, Stafford, Rodney Hopson and Henry Frierson. 2015. *Continuing the journey to reposition culture and cultural context in evaluation theory and practice*. Charlotte, NC: Information Age Publishing. (Evaluation and Society series).

- Hoppers, Catherine O., and Howard Richards. 2011. *Rethinking thinking: Modernity's 'other' and the transformation of the university*. Pretoria: University of South Africa. (Taproot Series I.)
- Ikuenobe, Polycarp. 2017. Tradition and a foundation for African renaissance. In: *Decolonising knowledge for Africa's renewal*, (ed.) Vuyisile Msila, 2–19. Randburg, South Africa: KR Publishing.
- ILO (International Labour Organization). 2020. COVID-19 and the world of work: Updated estimates and analysis. *ILO Monitor*, 4th ed.
- Khomba, J. K. 2011. Redesigning the balanced scorecard model: An African perspective. PhD Diss., University of Pretoria.
- Le Grange, Lesley. 2014. Currere's active force and the Africanisation of the university curriculum. *South African Journal of Higher Education* 28 (4): 1283–1294.
- Lent, J. R. 2017. *The patterning instinct: A cultural history of humanity's search for meaning*. New York: Prometheus Books.
- Maldonado-Torres, Nelson. 2007. On the coloniality of being: Contributions to the development of a concept. *Cultural Studies* 21 (2–3): 240–270.
- Mawere, Munyaradzi, and Gertjan Van Stam. 2016. Ubuntu/unhu as communal love: Critical reflections on the sociology of ubuntu and communal life in sub-Saharan Africa. In: *Violence, politics and conflict management in Africa: Envisioning transformation, peace and unity in the twenty-first century*, (eds.) Munyaradzi Mawere and Ngonidzashe Marongwe. Bamenda, Cameroon: Langaa RPCIG.
- Mertens, Donna. 2008. *Transformative research and evaluation*. New York: Guilford Press.
- Ministry of Environmental Affairs. 2008. Deputy Minister of Environmental Affairs, Ms Rejoice Mabudafhasi's speech on the occasion of Conservation with the People for the People Conference. Ministry of Environmental Affairs, Johannesburg.
- Mustonen, Tero, and Pauliina Feodoroff (with the Skolt Sámi Fishermen of Sevet-tjärvi). 2018. Skolt Sámi and Atlantic Salmon Collaborative management of Näätäimö Watershed. Finland as a case of Indigenous evaluation and knowledge in the Eurasian Arctic. *New Directions for Evaluation* 2018 (159): 107–119.
- Moyo, Dambisa. 2009. *Dead aid: Why aid is not working and how there is another way for Africa*. London: Penguin Books.
- Ndlovu-Gatsheni, Sabelo J. 2013. The entrapment of Africa within the global colonial matrices of power: Eurocentrism, coloniality, and de-imperialisation in the twenty-first century. *Journal of Developing Societies* 29 (4): 331–353.
- Noble, Safiya Umoja. 2018. *Algorithms of oppression: How search engines reinforce racism*. New York: NYU Press.
- Oelofsen, Rianna. 2015. Decolonisation of the African mind and intellectual landscape. *Phronimon* 16 (2): 130–146.

- O'Neil, Cathy. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Broadway Books.
- Oxfam. 2020. *Dignity not destitution: An 'economic rescue plan for all' to tackle the coronavirus crisis and rebuild a more equal world*. London: Oxfam.
- Patton, Michael Q. 2019. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Paipa, Kirimatao, Fiona Cram, Vivienne Kennedy et al. 2015. Culturally responsive methods for family centered evaluation. In: *Continuing the journey to reposition culture and cultural context in evaluation theory and practice*, (eds.) Stafford Hood, Rodney Hopson and Henry Frierson, 313–334. Charlotte, NC: Information Age Publishing (Evaluation and Society series).
- Picciotto, Robert. 2020. From disenchantment to renewal. *Evaluation* 26 (1): 49–60.
- Purcell, Sebastian. 2018. Life on the slippery Earth. <https://tinyurl.com/vwpdvbk>
- Rowe, A. 2019. Sustainability-ready evaluation: A call to action. In: *Evaluating sustainability: Evaluative support for managing processes in the public interest*, (ed.) G. Julnes, 29–48. Hanover, PA: Wiley Periodicals, Inc., and the American Evaluation Association.
- Raworth, Kate. 2017. *Doughnut economics: Seven ways to think like a 21st-century economist*. London: Penguin Random House.
- Russon, Craig. 2008. An Eastern paradigm of evaluation. *Journal of Multidisciplinary Evaluation* 5 (10): 71–77.
- Russon, Craig. 2014. Impact evaluation based on Buddhist principles. *Journal of Multidisciplinary Evaluation* 10 (23): 73–77.
- Russon, Craig, and Karen Russon. 2010. How the *I Ching* or *Book of Changes* can inform Western notions of theory of change. *Journal of Multidisciplinary Evaluation* 6 (13): 193–199.
- Schwandt, Thomas A. 2009. Globalising influences on the Western evaluation imaginary. In *The SAGE International Handbook of Educational Evaluation*, (eds.) Katherine E. Ryan and J. Bradley Cousins, 19–36. Thousand Oaks, CA: SAGE Publications Inc.
- Schwandt, Thomas A. 2019. Post-normal evaluation? *Evaluation* 25 (3): 317–329.
- Schwandt, Thomas A., and Emily Gates. 2016. What can evaluation do? An agenda for evaluation in service of an equitable society. In: *Evaluation for an Equitable Society*, (eds.) S. Donaldson and R. Picciotto. Charlotte, NC: Information Age Publishing.
- Stiglitz, Joseph E. 2002. *Globalisation and its discontents*. London: Penguin Books.
- Terreblanche, Sampie. 2014. *Western empires: Christianity and the inequalities between the rest and the West, 1500–2010*. Johannesburg: Penguin Books.
- Uitto, Juha I., Jyotsna Puri and Rob D. van den Berg (eds.). 2017. *Evaluating climate change action for sustainable development*. Springer Nature.

- UN (United Nations). 2009. Outcome document of the High-Level United Nations Conference on South-South Cooperation. A/64/L.37. Draft resolution submitted by the President of the UN General Assembly, 64th UNGA Session, 2009–2010. New York: United Nations.
- UN DESA (United Nations Department of Economic and Social Affairs). 2020. *The Sustainable Development Goals report*. New York: UN DESA.
- UNDP (United Nations Development Programme). 2020. COVID-19 and the SDGs. <https://feature.undp.org/covid-19-and-the-sdgs/>
- Waalaneekweew/Bowman-Farrell, Nicole R. 2018. Looking backwards but moving forward: Honoring the sacred and asserting the sovereign in Indigenous evaluation. *American Journal of Evaluation* 39 (4): 543–568.
- Waalaneekweew/Bowman-Farrell, Nicole R., and Carolee Dodge-Francis. 2018. Culturally responsive Indigenous evaluation and tribal governments: Understanding the relationship. *New Directions for Evaluation* 2018 (159): 17–31.
- World Bank. 2018. *The World Bank in China overview*. Washington, DC: World Bank.
- WEF (World Economic Forum). 2020a. *The global risks report 2020*, 15th ed. Cologny, Switzerland: World Economic Forum.
- WEF (World Economic Forum). 2020b. COVID-19: What are the biggest risks to society in the next 18 months? Blog post. 3 July. <https://tinyurl.com/6cp34w8y>
- Yanka, Dorji, Peter Newman, Vanessa Rauland et al. 2018. Sustainability in an emerging nation: Bhutan case study. *Sustainability* 10: 1622.
- Yunkaporta, Tyson. 2020. *Sand talk: How Indigenous thinking can save the world*, 1st ed. San Francisco, CA: HarperOne.

PART II

EXPERIENCES





CHAPTER 4

Ensuring Transformational Change for Climate Action

MATTHEW SAVAGE, TIM LARSON, JESSICA KYLE AND
SAM MCPHERSON

Abstract. This chapter sets out lessons learned and insights into transformational change arising from an evaluation of the Climate Investment Funds (CIF). It draws upon work undertaken during an independent evaluation of transformational change in the CIF during 2018 and 2019 (Itad 2019) and work that the evaluation team supported through the Transformational Change Learning Partnership (TCLP) after the evaluation. The CIF commissioned the evaluation to explore to what extent CIF had supported transformational change across a range of climate change areas: supporting clean energy and reducing greenhouse gas emissions (mitigation); reducing systemic risk and creating greater resilience to the impacts of climate change (adaptation); and enabling investments in sustainable forestry and strengthening the role of climate action in addressing other areas such as gender equity. We describe the baseline thinking on transformational change in the CIF that underpinned the Itad evaluation, describe findings that arose from the evaluation, provide insight into further work on transformational change that the Itad team undertook as part of the TCLP process and identify areas for further consideration and development. This article builds upon recent analysis of TCLP concepts and learning (e.g. CIF 2021, Williams, Dickman and Smurthwaite 2020).

Approach

The Climate Investment Funds (CIF) were established in 2008 to expand finance for climate change mitigation and resilience, filling urgent financing gaps and demonstrating the viability of emerging solutions. With more than \$8 billion contributed, CIF supports transformational change towards low-carbon, climate-resilient development in the areas of mitigation, resilience and forests through four programmes: 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). At the time of the evaluation, these programmes had collectively supported 300 projects in 72 countries.

The portfolio of CIF programmes is extremely broad in terms of its thematic and geographic coverage. Climate change is a cross-cutting topic that touches nearly every aspect of social, economic and environmental development. CIF projects range from global to local, cover a range of sectors (e.g. energy, transport, urban development, infrastructure, water, agriculture, forestry) and deploy a range of interventions, including technology, governance and capacity building, market creation, financing, behavioural change and policy development.

The CIF selected Itad Ltd., a strategic evaluation and learning consultancy, to undertake an independent evaluation of transformational change covering each of the four programmes. At the time of the evaluation, CIF had already set up the Transformational Change Learning Partnership (TCLP), bringing together a range of academics and practitioners to strengthen concepts and understanding of transformational change within the climate change sphere (CIF 2020a). Itad was invited to make further contributions to the existing TCLP frameworks to support the evaluation (CBI 2019). Initial work by Itad included a review of the concepts of transformational change developed under the TCLP¹.

- *Definition of transformational change.* The working definition of transformational change that the TCLP developed and the evaluation used was: 'Strategic changes in targeted markets and other systems with large-scale, sustainable impacts that accelerate

¹ As noted in the sections below, the initial TCLP frameworks were reviewed and advanced further in 2020 and 2021 and continue to evolve based on ongoing learning in the TCLP and stakeholder feedback (CIF 2021; Williams 2018; Williams, Dickman and Smurthwaite 2020).

or shift the trajectory towards low-carbon and climate-resilient development'. This formed the basis for the focus of the evaluation.

- *Dimensions of transformational change.* Transformational change dimensions are core characteristics for change to be considered on a path towards possible transformation in the context of climate change action. The TCLP had identified four dimensions that were incorporated into the evaluation – *relevance, systemic change, scale and sustainability* (box 4.1). These four dimensions collectively captured elements of transformational change that ought to be present. The TCLP recognized that, although these dimensions might vary in emphasis and significance (based on context and timing), all must be attended to or present to some extent for there to be confidence that climate change actions are relevant to transformational change.
- *Arenas of transformational change.* The arenas of Intervention had been developed during an earlier portfolio review of CIF-supported programmes and projects (Ross Strategic and Community Science 2017). Arenas describe the types of interventions commonly made within CIF programming to advance climate action (table 4.1). Actions within and between these arenas can be designed and implemented to advance systemic changes, to expand pathways, to enhance the sustainability of changes, to speed progress and to increase the relevance of changes to goals or contextual factors.

Box 4.1 Transformational Change Learning Partnership: Definitions of Transformational Change

- *Relevance:* Strategic focus, design and nimbleness of initiatives to enable transformation
- *Systemic change:* Fundamental shifts in system structures and functions
- *Scale:* Contextually large-scale transformational processes and impacts
- *Sustainability:* Robustness and resilience of changes

Signals of Transformational Change

Having refined the definitions and dimensions, the evaluation team faced a more practical challenge – how to capture evidence of transformational change in practice in a way that could support a robust evaluation. The

Table 4.1 Arenas of Transformational Change

Arena of intervention	Definition
Financing	Interventions that leverage, complement and coordinate other funding sources to develop financing structures over time, with a focus on crowding in private sector financing. Interventions that use capital to buy down costs or cover risks in ways that lower longer-term costs and risks through economies of scale and market transparency and development and use financial incentives to shift behaviours and decisions in ways that accelerate deployment of low-carbon and climate-resilient development.
Governance and engagement	Interventions that build strong, durable country ownership and support for CIF-supported interventions; ensure meaningful inclusion, engagement and empowerment of relevant parties (including women and indigenous peoples) or ensure that the full range of salient barriers to transformation are identified and addressed using a programmatic approach.
Institutions	Interventions that focus on building or strengthening institutional capacity of key public sector (national, regional, local) and civil society organizations operating within the country. Interventions that develop or enhance institutional communication, coordination and collaboration among organizations working in the country, including multilateral development banks and other international partners.
Knowledge and information	Interventions that generate, share or diffuse information to enhance knowledge and expertise to support accelerated implementation of low-carbon and climate-resilient development, including research and analysis, measurement and evaluation, learning partnerships, and training and capacity building for local populations.
Markets	Interventions that expand private sector awareness, capacity and opportunities to enter and successfully participate in markets that advance low-carbon and climate-resilient development, such as renewable energy technologies, low-carbon transportation, sustainable forestry and ecosystem services. Interventions that establish clear, predictable market rules, mechanisms, relationships and infrastructure to overcome barriers and support private-sector market involvement.

(continued)

Table 4.1 Arenas of Transformational Change *(continued)*

Arena of intervention	Definition
Natural capital	Interventions that work with natural systems to reduce greenhouse gas emissions or make other physical changes to increase ecosystem resilience, including reforestation and enhancement of forest carbon stocks; increasing the agro-ecological potential of an area; enhancing blue carbon attributes of aquatic and coastal ecosystems and restoring habitat to protect native species, preserve biodiversity or improve ecosystem health.
Policies	Interventions that support development or testing of laws, policies or regulations that create an effective enabling environment for deploying low-carbon and climate-resilient development solutions, including laws and regulations promulgated through formal legislative and public sector policy-making processes – as well as through policies and plans – and established by key institutions.
Practices and mindsets	Interventions that seek to influence individual or private sector practices, decisions and behaviours using tools and techniques drawn from social marketing and other fields, often involving shifting mindsets and individual-level appreciation of opportunities and benefits and recognizing the power of social bonds and relationships in establishing and reinforcing norms and practices.
Technologies and infrastructure	Interventions that support first use of key technologies in a country to demonstrate their effectiveness, develop technology deployment competencies in the private and public sectors and drive reductions in technology deployment costs and risks (e.g. through economies of scale, implementation data to inform investment risk assessments). Interventions that improve the infrastructure necessary for low-carbon and climate-resilient development.

Source: Ross Strategic and Community Science (2017).

TCLP's working definition and four dimensions of transformational change provided a starting point, but they lacked the granularity, forward-looking perspective and dynamism necessary to recognize transformational change in different contexts, at different country and geographical levels, in different sectors and at different timescales.

There was an obvious need to make the framework more practical to support collation and analysis of quantitative and qualitative evidence

relating to the CIF programmes, but many CIF stakeholders, although comfortable with the high-level transformational change definition and dimensions, struggled to articulate what these might look like in practice. A common refrain was that CIF stakeholders 'would know transformation when they saw it' but were less comfortable in creating specific transformation indicators or benchmarks against which programmes might be assessed or measured.

Discussions on trying to create a coherent framework identified a number of challenges, which reflected the complex nature of transformational change itself (Itad 2019; Williams 2018; Williams, Dickman and Smurthwaite 2020):

- Transformation can refer to changes at different scales, from individual to global.
- Transformation can appear in many different forms depending on the sector and context.
- Data on programmatic transformational change are generally weak from a monitoring and evaluation viewpoint.
- Transformation occurs relative to dynamic baselines, which are often poorly documented.
- Transformation involves addressing multiple barriers or constraints in parallel.
- Transformation usually occurs beyond programme boundaries, where results chains are weak.
- Timescales of transformation are typically longer than those of supporting projects.
- Transformation is dynamic and non-linear and requires sequential, multistage interventions.
- Transformation, as a complex system change, can be influenced but not controlled.

Recognizing the need for an innovative approach, the evaluation team considered the work of other institutions grappling with similar challenges, including the World Bank's review of transformational engagements (World Bank Group 2016), the experience of the U.K. International Climate Fund in developing its key performance indicator on transformational change (DFID 2014) and the Initiative for Climate Action Transparency's draft guidance on transformational change evaluation (ICAT 2020). A review of these efforts, coupled with the TCLP's work to date, provided additional thematic insights:

- Transformation often emerges with a sequential pattern as part of a process over time.
- Signals of transformation are found in outcomes and processes that support them.
- Signals of transformation can be broadly mapped using the dimensions of transformation.
- The long-term nature of transformational change requires proxies to capture the likelihood of future change.

Based on these insights, the team created a framework centred on the concept of *signals of transformational change*. Signals were defined as system characteristics that demonstrate progress towards transformation, whether at early, interim or advanced stages. The team consciously developed the framework to be indicative, rather than prescriptive, with signals based on qualitative and descriptive information in addition to quantitative data. The signals also included proxies for future change that might not be quantifiable during or immediately after project or programme implementation. The resulting signals framework considered captured three simultaneous aspects of transformational change: stages, dimensions and sector or theme (CIF 2020b).

Signals Over Time

Signals of transformation typically emerge and strengthen over time, often over the course of years, starting with early signals based on programme design and extending to long-term outcomes after programme completion. Although progress is not always linear, stages generally follow a pattern. Three stages were identified for the evaluation.

- *Early signals*. Relevant programme design and implementation are enabling preconditions for transformation.
- *Interim signals*. Interim outcomes external to the programme boundaries are evident. This includes process advancements such as policy development and budget allocation that support and advance progress towards transformational outcomes over time.
- *Advanced signals*. Long-term, self-sustaining outcomes are materializing.

The context in which change occurs and the ambition of the transformational change are worth noting in relation to stage of advancement. What

might be regarded as modest capacity advancements in a developed market or governance context might be more fundamentally transformational in a less-developed country context; therefore, framing around advancements should be considered in context. Progress is also not always assured or linear. Setbacks can occur, and context, such as local resource availability, can change, making earlier progress less relevant. For example, cost reductions associated with one type of renewable energy source (e.g. photovoltaic solar power) may outpace cost reductions associated with another renewable energy source (e.g. geothermal power). In this case, advancement can slow or even come to a halt in less cost-competitive technology markets and pick up speed in more cost-competitive markets. Similarly, extreme weather events, political upheaval, global economic downturns and other events can slow or reverse progress on climate action in uncontrollable ways. For example, the Arab Spring in 2010 affected the CIF's attempts to establish a concentrated solar power (CSP) programme across the Middle East and North Africa. For these reasons, advancement in a linear and predictable fashion is not assumed, and the ability to be nimble and adapt design, strategy and implementation are paramount to ultimate success.

Signals Across Dimensions

The team recognized that signals of transformation could be mapped broadly against the four dimensions (relevance, systemic change, scaling, sustainability). Relevance was considered an early signal, as set out above, that programmes had been designed for transformational success, although relevance should not be taken for granted, and programme design often needs to be revisited over time to remain relevant. Scaling and sustainability are likely to emerge in the longer term.

A framework that captures these signals and was used to inform the evaluation is set out in table 4.2.

Signals Across Sectors and Thematic Areas

Although some signals are universal to all types of development programming (e.g. capacity development), many signals differ substantially according to sector or thematic area. For example, progress towards climate-resilient agriculture differs from progress towards utility-scale grid decarbonization. Based on the evaluation case studies, the evaluation team compiled illustrative signals according to sector or theme in addition to stage and dimension. These sector- and theme-based signals focus on the interim

Table 4.2 Signals of Transformational Change in the Climate Investment Funds

	Interim signal	Advanced signal
Systemic change	<i>Meaningful progress on activities to overcome barriers (e.g. new institutions and capacity, enhanced governance structures, new policies and regulations, new planning processes, new financing structures)</i>	<i>Evidence of system change outcomes that influence decisions or behaviours (e.g. changes in planning decisions and outcomes, uptake of incentives, changes in budgetary allocations, increased awareness, changes in consumption or access patterns, greater affordability, greater technology availability)</i>
Scale	<i>Increased activity that might facilitate scaling (e.g. new finance programmes and investors, evidence of pipeline development, supply chain expansion, new distribution networks, new access and delivery platforms)</i>	<i>Evidence of scaling outcomes (e.g. more market participants, increasing financing flows, large-scale greenhouse gas emission reductions, number of consumers and service users, increased sales of new technologies, increased geographic coverage, increased national-subnational linkages, increased community participation and uptake)</i>

Source: Adapted from Williams, Dickman and Smurthwaite (2020).

and advanced stages, given that early-stage signals are more generic. The team therefore developed specific signal frameworks for each programme (low-carbon infrastructure, energy access, adaptation, forestry).

Analysis

The Itad team successfully applied the transformational change framework and used it to identify and assess progress within the CIF. Each of the four major CIF programmes was analysed against the dimensions, as well as the strength of signals within them (early, interim, advanced). Progress on transformation was much more robust in the large, low-carbon CTF programme than in the other programmes (SREP, PPCR, FIP), in part because of its geographic focus on middle-income countries with greater capacity and in part because

of its thematic focus on energy deployment, a sector in which progress on energy technology and innovation has been more robust than in other climate sectors. Other programmes were dealing with poorer or lower-capacity countries or were seeking to transform much deeper and more complex social and environmental systems, requiring longer-term engagement.

At a higher level, the evaluation identified a number of lessons in the CIF that are relevant to the broader understanding of transformational change²:

1. Signals of transformational change emerge in at least a partially sequential manner over time.

The evaluation identified that there is a flow through the dimensions that can, to some extent, mirror the early, interim and advanced framework. For example, signals in the relevance dimension are associated with the design and implementation phase and are correlated with early signals in terms of their maturity (creating conditions for change), noting that programmes must revisit their mandates and designs over time to ensure that they remain relevant. Systemic change and scaling signals tend to arise towards the end of and after project implementation, with signals of sustainability emerging later as the resilience and robustness of other dimensions are tested. It was therefore not surprising that the evaluation found more-advanced signals of relevance and only earlier signals of sustainability for three of the four CIF programmes, particularly because many country programmes (particularly SREP and FIP) remained in early implementation.

2. Two basic transformational models of transformational change were identified in CIF programming: scale to systems and systems to scale.

The evaluation had shown that early signals of the impacts of transformational change may be modest or even barely discernible because a sufficient number of *systemic changes* is needed to overcome barriers and foster enabling conditions that enable later accelerated *scaling*. In other cases, early scaling of a change through large-scale investment can catalyse systemic changes that can in turn create a feedback loop for further scaling. In both cases, there can be dynamic interplay between systemic change and scaling. These two transformational change models were identified

² These build on similar lessons and reflections that have been covered in other publications (e.g. Van den Berg, Magro and Salinas Mulder 2019; Williams 2018; Williams, Dickman and Smurthwaite 2020).

within the CIF portfolio, recognizing that, at times, transformation can be advanced through a more-simultaneous mixture of the two models, as well as through other means.

- The first model (more prevalent in the largest CTF programme) uses a scaling-based approach, deploying large volumes of concessional finance to demonstrate the feasibility of new approaches or technologies, reduce investor and policymaker perceptions of risk, increase transparency regarding costs and operational performance, and reduce the costs of delivery (through economies of scale). These were typically large investments in utility scale generation (solar photovoltaic, wind, geothermal, CSP). Project sizes were typically in the hundreds of millions and sometimes billions of dollars. It was expected that systemic change and further replication would follow as policymakers, developers and investors adjusted their risk perceptions and mobilized further large-scale finance. Sustainability is achieved through subsequent adjustments in the policy environment and sustained investor interest.
- The second model (more prevalent in SREP, PPCR and FIP) is delivered through a systems-change lens. It is structured around capacity building, awareness raising, strengthening the enabling environment, institutional strengthening and governance, and piloting of smaller-scale interventions to deliver proof of concept. It is hoped that, by improving the underlying system, scaling then follows as the enabling environment becomes more supportive of change, pilot projects prove successful and other investors and project developers choose to move into the investment space. The focus of this model may depend on the stage of market development, with low-income countries requiring more attention to awareness, capacity and governance and middle-income countries more oriented towards private sector incentives, risk reduction and competitiveness.

3. Transformation is more likely to occur quickly when a broad range of project outcomes and contextual factors align, making transformational change a dynamic, unpredictable process.

Transformational change requires alignment of a range of factors, some of which are project related and others of which occur in the external contextual environment. For example, in clean energy markets, influencing

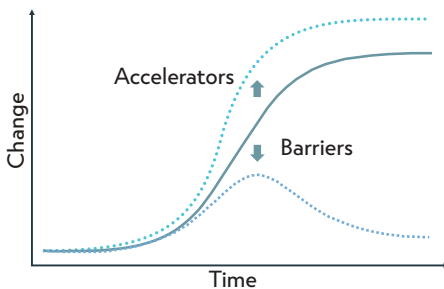
factors supporting transformation have included a facilitating regulatory environment, a robust investment climate, access to affordable finance, an increase in the availability of cost-competitive technologies, strong consumer awareness and demand, and clear political will to shift towards a clean development trajectory. The absence of a single element can lead to delayed take-off, with transformation not becoming apparent until after several years of modest results or not occurring at all. Transformation appears to occur more quickly in middle-income countries with stronger enabling environments and markets that are closer to tipping points, with examples of countries leapfrogging to bypass existing support mechanisms, than in less-developed markets (with lower capacity and financing constraints) or more-contested sectors (e.g. forestry and community-level resilience), where timescales for transformation can be much longer.

4. Incremental changes make a valuable contribution to progressing towards future transformation but are not in themselves transformational.

Given the timescales and uncertainty associated with transformation, incremental change is important in terms of laying the groundwork for future change and potential tipping points. The evaluation suggested that activities such as capacity building, changing mindsets and altering behaviours can have a cumulative transformational effect, the results of which become clear only when change processes that rely on these foundations later occur. Incremental change will often be the most likely pathway for a time-limited programme when there are significant weaknesses in the operating environment

(e.g. development challenges, political instability, resource constraints) or technologies remain far from commercialization. Realistic expectations are therefore required regarding the likelihood of transformation during programme implementation cycles. Nevertheless, incremental change and reform are not the same as transformation, although they may lead to transformation in some contexts. Figure 4.1 shows how some changes may help accelerate transformational

Figure 4.1 Transformation Pathways Under Different System Contexts



Source: Developed by authors as part of the TCLP process.

processes, whereas others may be insufficient to overcome barriers that enable transformational processes to unfold and take off.

5. Transformational change typically involves shifts in power.

Shifts in power, decision-making authority, inclusion and distributional effects of change are common in climate change transformations and can occur as part of systemic changes that create the enabling conditions for change. They can also occur as change scales and the distributional effects of large-scale change alter the locus of economic and political power. Power shifts can manifest between institutions (e.g. energy and environment ministries), levels of government and private sector actors and along other axes. Resistance to power shifts can increase barriers to transformation, whereas expanded access to power can have a snowballing effect that accelerates transformational processes. Such shifts in power can play out in disruptive or smooth ways, depending on context and the characteristics of change.

6. The timescales of transformation processes must be acknowledged, and the assessment of transformation must be assessed relative to context and opportunity.

The evaluation sought to classify evidence in terms of stage of transformation (early, interim, advanced). There were challenges to this in that such categorization appeared to offer a potentially negative value judgement on the performance of programmes identified as being in the early or interim stage. Programme managers were sensitive to the accusation that their programmes had in some way failed to be transformational. It is important when deploying these frameworks to recognize that some programmes (particularly those that adopt a systemic change model) may deliver transformation over long time horizons. The temporal or process element should therefore be non-judgemental and simply seek to capture the stage that the transformation process has reached.

7. A portfolio approach offers a balance of short- and long-term transformational change programmes, focusing on pathways relevant to different sectors and contexts in appropriate ways.

The CIF portfolio supported a range of projects, some of which reached tipping points (with scaling and sustainability likely in the short term) and some of which prepared the ground for much longer-term systemic change. Although there may be some value in prioritizing scarce resources

towards early action (from greenhouse gas mitigation and climate adaptation perspectives), this should not be at the expense of projects that are equally important over the medium to long term but may face greater challenges, whether from a technology, sector or country-context perspective. Investing in such projects creates an options value for larger-scale future transformation. A broad climate finance portfolio such as CIF also allows winners and best practices to emerge and can generate lessons that may be fed back into other projects. For example, it is not clear whether CSP will emerge as a competitive technology versus the improving economics of solar photovoltaic plus battery storage as a solution to providing dispatchable power. From this perspective, there is value in ensuring good portfolio diversification (e.g. across themes, country contexts and technologies) and using learning for course correction and improved programming.

8. Transformational change occurs in complex environments, and evaluation focus should be on establishing contribution rather than attribution.

Programmes and projects can contribute to transformational change, but there are often many other actors, initiatives and forces at play. Multiple influences shape how complex systems evolve: sometimes in aligned directions, sometimes in quite different directions. At the same time, events and trends unfold that shape the context for change in evolving and disruptive ways. The ability of a programme or project to catalyse, contribute to or support shifts and transformation in a complex system is often mediated through this larger dynamic context of activities, actors and forces. This reality often creates challenges for clearly assessing the contribution of individual programmes, projects and actors to transformational change. Evaluations should therefore seek to demonstrate the contribution case for individual programmes rather than to establish attribution.

Impact of the Evaluation

The evaluation of transformational change of the CIF was well received for the objectivity and usefulness of its findings and for the contribution it made to helping advance thinking regarding concepts of transformational change.

The evaluation was an important milestone document that was able to provide evidence to validate the CIF programmatic model in a number of

ways, recognizing its uniqueness among global climate funds (Itad 2019). Identified CIF strengths included:

- A programmatic approach built on investment planning processes with governments and a range of other stakeholders
- The predictability and flexibility of large-scale funding provided by CIF programmes
- Coordination and alignment of multiple multilateral development banks around national objectives
- Mobilization of key political champions and change agents for implementation

The evaluation was presented as a core part of the CIF 10-year anniversary meeting at Ouarzazate, Morocco, and its findings formed the basis for subsequent CIF Trust Fund Committee discussions that resulted in the decision to continue the CIF and further develop its offering, including substantial financial replenishment and the launch of new thematic programmes.

The evaluation findings and the conceptual frameworks for transformational change analysis were also taken forward in a number of case studies, including on country programmes (e.g. Zambia resilience programming under the PPCR) (CIF 2020c) and thematic areas (e.g. CTF support for CSP (CIF forthcoming)). In both cases, the evidence gathered during the evaluation was presented in more depth than in the overall evaluation, and further stakeholder discussions were held to explore how transformation could be better reflected in national policy and CIF programmatic design.

Subsequent Development Under the TCLP

Having completed the evaluation, several Itad team members have continued to engage on the topic of transformational change through the TCLP process and have developed additional theoretical frameworks to strengthen and deepen understanding of transformation processes in climate action. Subsequent developments include a number of focus areas.

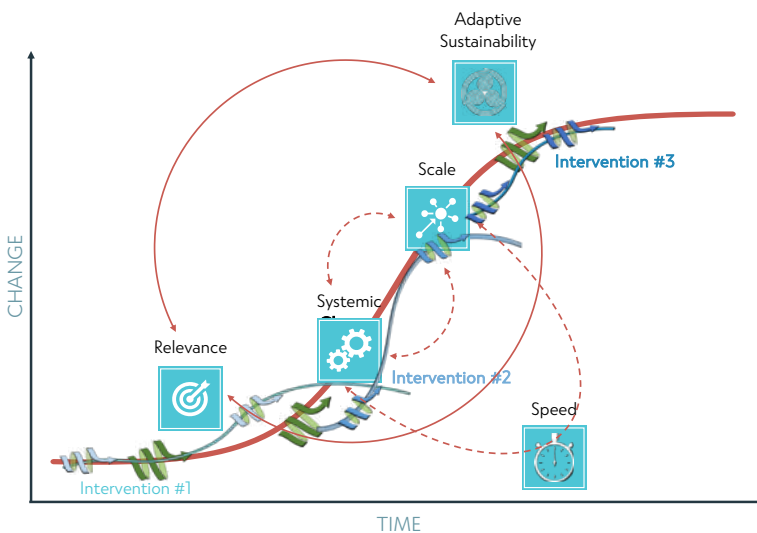
First, team members began to review the linkages between the dimensions to look for patterns and relationships reflected in observed transformation processes. Although transformational change in complex systems often unfolds in winding, convoluted, unpredictable ways, patterns relevant to adoption and diffusion of specific actions, technologies and

practices can be discerned in transforming systems. A legacy of studies on the diffusion of innovation, technologies and practices indicates the potential usefulness of the classic s-curve for understanding and thinking about the diffusion or adoption of climate actions (figure 4.2).

The s-curve diagram shows that change does not happen in a linear way, although it shows how progress in diffusing and scaling climate actions may lag because systemic changes and other groundwork is needed to foster the enabling conditions and overcome barriers that enable change. In this stylized s-curve model, there is dynamic interplay between the dimensions of change. Transformational processes and diffusion and adoption of climate actions can vary widely in the curves they actually follow.

Second, team members, in discussions with the CIF Evaluation and Learning Initiative and TCLP members, came to recognize the need for a new dimension of transformational change – speed. Transformation takes place over different time frames and at different speeds (figure 4.3). Substantial work on systemic changes that create preconditions for transformation may not manifest in clear results for some time but may be followed by significant scaling and impact. Change processes are not linear and often happen in fits and starts, sometimes with backsliding and sometimes with rapid acceleration and scaling. The speed dimension captures

Figure 4.2 Stylized S-Curve Model of Diffusion of Climate Actions in Transforming Systems

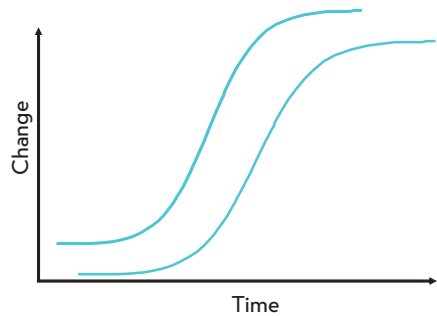


Source: Developed by authors as part of the TCLP process.

evidence of the timeliness of transformational change processes and outcomes and their temporal alignment with desired transformation pathways. In a programme life cycle, speed signals can be observed in the design phase (e.g. considerations of timing and acceleration), during implementation (e.g. ensuring appropriately timed actions and outputs that support delivery) or after the programme is finished (delivering outcomes and impacts that reflect the necessary pace of transformation).

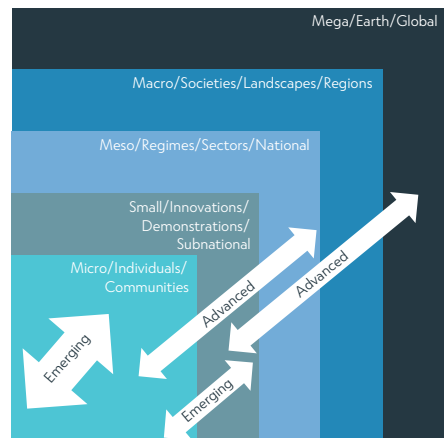
Third, transformation must occur across different system scales to be meaningful. Transformation at the level of depth and breadth needed to address the climate crisis is an extremely ambitious global goal, requiring changes spanning natural and human systems. Changes relevant to this transformation must occur at many levels (macro, meso, micro), although the larger scale ultimately matters for climate action. Positive transformations supporting climate action can happen in households and communities and at other levels. Although these changes can be valuable and beneficial on their own, the urgency of the climate crisis necessitates expansion of changes to higher systems (national, global). Likewise, changes that occur at higher levels, such as policies and regulations, must be fully mainstreamed at lower levels (subnational, local, individual) for them to be truly effective and embedded. Being able to link the different levels as part of transformational change processes is therefore key, with advanced change occurring only when higher- and lower-level systems are connected (figure 4.4).

Figure 4.3 Bringing Forward Transformational Change Pathways in Time



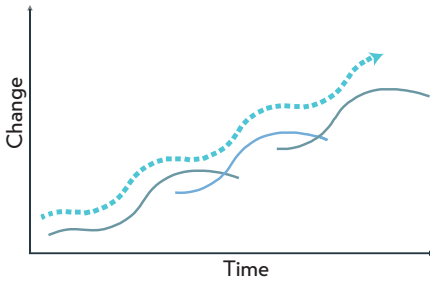
Source: Developed by authors as part of the TCLP process.

Figure 4.4 Multilevel Perspective on Transformational Change



Source: Developed by authors as part of the TCLP process.

Figure 4.5 Multi-Stage Approach to Transformation



Source: Developed by authors as part of the TCLP process.

Finally, sustainability is not the same as system stability. As systems are transformed, new equilibriums emerge in which the systemic changes support a 'new normal' of decisions, actions and practices. The system reaches a point at which the old paradigm has been sufficiently displaced and the likelihood of backsliding or regression to the former state becomes unlikely, although sustainability should not be viewed as a stable state or 'final destination'. Systems continue to be

subject to emerging transformational pressures and dynamics and adapt accordingly. Different technologies and market solutions may emerge and compete for dominance over time. Figure 4.5 shows how successive waves of transformational change can build over time as systems adapt and evolve. Programmes therefore need to be agile, nimble and adaptable, even if overall goals remain the same. By renaming sustainability 'adaptive sustainability', we integrate the above concepts.

Areas for Further Development

There are a number of areas for further exploration that arise out of the evaluation and further work within the TCLP.

Definitions of Transformational Change

One of the challenges has been to create a definition of transformational change that can capture high-level systemic change while being useful for individual practitioners in the field of climate change. To be credible, definitions must not only provide a global conceptual framework, but also facilitate practical application of this framework by those designing programmes, projects and interventions. Broader usability of the transformational change framework requires further consideration and development, including additional guidance and examples of real-world applications.

Dimensions of Transformational Change

The dimensions have proved useful for identifying the elements of transformational change (e.g. within the s-curve), as well as providing a useful framework on which to categorize signals, although there continue to be challenges in ensuring that the dimensions are clear and capture the full range of elements. Of particular interest are the relevance and adaptive sustainability dimensions, whose definitions and boundaries between them have proved challenging. Speed also presents challenges as a concept, incorporating aspects of appropriate timing and acceleration. Finally, there are potential definitional boundary issues between systemic change and scaling (where the concept of scale is implicit in changing system function). These are set out in more detail here.

- *Relevance.* This dimension was used in the evaluation context as a starting point for transformational assessment – is this the right approach to the right problem at the right time, and are the conditions for transformation mainstreamed into the programme? It has become clear that relevance contains a directional or normative element (Is this the right direction of travel?), as well as a practical element (Is this the right intervention to get us there?). It is also clear that relevance is present throughout the transformational change process, in that there is an ongoing need to constantly review the direction, assess any changes in the contextual environment (political, technological, social) and be prepared to adjust course (or potentially let interventions go when they are no longer contributing).
- *Sustainability.* Sustainability has been a challenging concept for a number of reasons. First, it suggests a somewhat static state that may continue in perpetuity and fails to capture the dynamic evolution of systems and markets over time. This requires some acknowledgement of the adaptive nature over time. Second, the term itself can be misinterpreted as relating to environmental impacts, rather than to the robustness or resilience of outcomes, to the notion of dynamic equilibrium. Although equilibrium may imply a level of alignment with normative views on environmental, social or economic sustainability, no single one of these should be the sole determinant.
- *Speed.* Although the team has acknowledged that speed is a new and important dimension (compressing the timescales along the

x-axis of the s-curve model and steepening the gradient), some concerns remain that there is also an element of timeliness that must be taken into account. Not all innovations or transformations can be achieved over compressed periods of time, and changes should not be forced when the contextual environment (e.g. technology availability, cost, social acceptance) does not support this.

- **Arenas.** Although the arenas represent barriers and opportunities within the systemic change dimension, they are also a useful lens for analysing the scaling and adaptive sustainability dimensions. Although they provide a comprehensive framework, the boundaries between the arenas are often blurred, with some level of overlap. Initial work has therefore focused on grouping of arenas into three higher-level categories: techno-economic, socio-institutional and environmental. Collectively, these provide an overarching framework on which signals can be organized under each dimension (with the exception of speed).

Signals of Transformational Change

The signals framework provided a useful starting point for the evaluation in capturing and categorizing evidence of transformation against the dimensions and a temporal or process axis, although this could be made more robust. Further work is ongoing under the TCLP in this regard that will include simplifying the stages of transformation and differentiating signals that indicate processes and progress towards transformation (emerging) from those that capture macro-level systemic shifts and alignment (advanced). As part of the advanced category, attention is being drawn to ensuring signals that address multiple levels of systems (micro, meso, macro), because without clear linkages and alignment, transformation is unlikely. A need has also been identified for a way to capture signals that indicate the absence of transformation (e.g. evidence of reinforcement of business-as-usual pathways) or even negative dynamics (backsliding or regression). A series of questions and indications of progress on moving from emerging to advanced signals is being explored to better capture and communicate progress across the dimensions and on transformational change overall. More-detailed sectoral sets of signals and guidance are also being developed, along with use cases to communicate to practitioners how these might be applied.

Acknowledgements

The authors would like to acknowledge the collective work of the CIF Evaluation and Learning Initiative team and the TCLP participants who have contributed to the conceptual framing, learning and insights of the evaluation and the subsequent development of additional conceptual frameworks. The authors recognize the inputs and feedback provided by Joseph Dickman, Senior Evaluation and Learning Specialist at the Climate Investment Funds.

References

- CBI (Consensus Building Institute). 2019. Climate investment funds: Transformational change learning partnership. Reflections report. <https://tinyurl.com/ydk237k6>
- CIF (Climate Investment Funds). 2020a. Climate Investment Funds: The Transformational Change Learning Partnership. <https://tinyurl.com/b2xjcbj6>
- CIF (Climate Investment Funds). 2020b. Signals of transformational change: Insights from the evaluation of transformational change in the CIF. <https://tinyurl.com/43afxwc>
- CIF (Climate Investment Funds). 2020c. Transformational change case study: Zambia: Building a resilient future. <https://tinyurl.com/zrrr2es3>
- CIF (Climate Investment Funds). 2021. Transformational change concepts. <https://tinyurl.com/xf5jh8r>
- CIF (Climate Investment Funds). Forthcoming. Transformational change in CSP.
- DFID (UK Department for International Development). 2014. ICF KPI 15: Extent to which ICF intervention is likely to have a transformational impact. DFID.
- ICAT (Initiative for Climate Action Transparency). 2020. Transformational change methodology: Assessing the transformational impacts of policies and actions. United Nations Environment Programme DTU Partnership and World Resources Institute.
- Itad. 2019. Evaluation of transformational change in the Climate Investment Funds. <https://tinyurl.com/yphmtcwp>
- Ross Strategic and Community Science. 2017. Climate Investment Funds: Synthesis of transformational change concepts. Available upon request. Contact ciftclp@worldbank.org.
- Van den Berg, Rob D., Cristina Magro and Silvia Salinas Mulder (eds.) 2019. *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*. Exeter, UK: IDEAS.
- Williams, Anna. 2018. CIF and transformational change: Essential learning for climate action. <https://tinyurl.com/4fkjd45n>

- Williams, Anna, Joseph Dickman and Regan Smurthwaite. 2020. Advancing evaluation and learning on transformational change: Lessons from the Climate Investment Funds' Transformational Change Learning Partnership. *American Journal of Evaluation* 42 (1). <https://tinyurl.com/y2x9tjuh>
- World Bank Group. 2016. *Supporting transformational change for poverty reduction and shared prosperity: Lessons from World Bank Group experience*. Washington, DC: World Bank Group.



CHAPTER 5

Evaluation at the Nexus Between Nature and Humanity for Transformational Change

JUHA I. UITTO

Abstract. The global environmental crises being manifested through climate change and rapid loss of biodiversity require transformational change in major systems ranging from energy and transportation to agriculture and cities. The pandemic of 2020–21 has demonstrated the interdependence of human and ecosystem health. Evaluation can contribute significantly to identifying solutions for the future but, to do so, must rise above its focus on individual interventions in isolation of their context. Evaluators must also learn to operate in the nexus between human and natural systems, where sustainable development takes place. This chapter draws upon experiences with evaluating the work of the Global Environment Facility (GEF) in supporting adaptation to climate change, an area that by necessity transcends the boundaries of human and natural systems. The chapter also introduces a framework for evaluating the GEF's additionality in six specific areas: environmental, legal and regulatory, institutional and governance, financial, socioeconomic and innovation.

Introduction

The state of the global environment and climate change have emerged, in the words of United Nations Secretary-General António Guterres, as the defining challenge facing humankind in the 21st century (United Nations 2018). The changing climate poses a long-term threat to the natural environment and human welfare. Its consequences are already being felt around the world in increasing weather anomalies, such as increasing frequency and intensity of storms, heatwaves and wildfires that directly affect how we live our lives and how our economies develop; climate change is also associated with societal conflict at many levels (Burke, Hsiang and Miguel 2015). Even if greenhouse gas emissions stopped tomorrow – an obviously impossible scenario – the lengthy lifetimes of such gases in the atmosphere would guarantee continued warming for several decades to come. There is thus an urgent need to enhance our capacity to adapt to climate change. This does not mean that we should give up on mitigation efforts. On the contrary, these need to be intensified to avoid catastrophe. At the same time, chemical pollution is reaching critical levels, posing great risks to human and ecosystem health. Business as usual will not do the trick; we need long-term transformations in our industrial, energy, urban, agricultural, transportation and other major systems to address the climate challenge.

The climate crisis, unfortunately, is not the only environmental challenge that we face. Biodiversity – species of animals and plants as well as entire ecosystems – is being lost at a faster rate than ever during the existence of humans on the planet (Caballos, Ehrlich and Dirzo 2017). Biodiversity, and all life on Earth, has intrinsic value in itself, but we are also losing resources that are very valuable to humans and society in terms of ecosystem services, including clean air and water, protection against storm surges and rising sea levels and medicines. Undisturbed ecosystems tend to find an equilibrium that benefits a multitude of species, which the reintroduction of wolves, an apex predator, to Yellowstone National Park has famously demonstrated, triggering a tritrophic cascade that has led to healthier populations not only of animals, but also of plants (Ripple and Beschta 2012).

Scientists are also increasingly realizing that the escalating outbreaks of new diseases, such as COVID-19, which became a global pandemic in spring 2020 with devastating human and economic costs, are directly linked to how we interact with and abuse the natural environment (UNEP 2016). Such zoonoses, in which a pathogen spills over from a non-human host to humans, cause 60 per cent of all infectious diseases and 75 per cent of emerging infectious diseases (Asokan and Asokan 2015). Interaction between species

requires adoption of a One Health approach, recognizing that ecosystem and human health are closely interlinked. As humans encroach deeper into ecosystems – building roads, clearing forests, mining – we disrupt ecosystems and come in ever-closer contact with wildlife, which makes it easy for pathogens to cross over to humans from non-human animals. Human population growth, unchecked urbanization and suburbanization and pursuit of financial profits drive these processes.

The good news is that environmental challenges are receiving more international attention than ever before. The United Nations member states adopted the 2030 Agenda for Sustainable Development and the attendant Sustainable Development Goals (SDGs) in 2015 (UN DESA n.d.). Environment is recognized as one of the three main pillars of sustainable development and can be seen as the foundation on which social and economic development depends. The Paris Agreement (UNFCCC 2021), through which an overwhelming majority of the world's countries agreed to limit their greenhouse gas emissions, came into being in the same year. The Global Commission on Adaptation (GCA 2021) (led by former United Nations Secretary-General Ban Ki-moon, International Monetary Fund Executive Director Kristalina Georgieva and Bill Gates) released its landmark report *Adapt Now* in September 2019, making an urgent call for leadership on climate resilience (GCA 2019).

Despite the new institutional arrangements and agreements, the steps that have been taken are not enough to halt climate change or species extinction or for the world to reach sustainable development, as the 2018 special report of the Intergovernmental Panel on Climate Change made clear (IPCC 2018). We need concerted efforts to address these challenges while the world is facing increasing uncertainty, and suspicions between countries and groups are on the rise. Financial resources, especially from public sources, are significantly smaller than the economic forces that contribute to climate change (e.g. fossil fuel subsidies, agricultural practices that lead to deforestation), although the role of public finance in promoting adoption and development of climate-friendly technologies is crucial (Van den Berg and Cando-Noordhuizen 2017). According to the Climate Policy Initiative, total annual flows of climate finance from public and private sources rose to \$590 billion in 2017/18 (Buchner et al. 2019). Financing, often in the form of subsidies, dwarfs these financial flows, to the detriment of the environment. We therefore must make sure that efforts to solving these challenges are effective and making a difference in the real world. This is where evaluation comes in. Evaluation can play an important role among other tools to provide evidence of the effectiveness, efficiency and impact of the various

policies, strategies, programmes and projects for transformational change. To do this meaningfully, evaluation must rise above tracking the results of individual initiatives in isolation and focus on what has been called 'significance' (Feinstein 2019) or 'transformational fidelity' (Patton 2020a).

In this chapter, I discuss the implications for evaluation, making a strong case that evaluation must consider all interventions in their broader context and how they interact with human and natural systems. It is not adequate to evaluate an intervention only against its internal logic without considering how it interacts with the external system, as well as any unintended consequences it may have (see e.g. Patton 2020b; Garcia and Zazueta 2015). I draw on evaluations by the Independent Evaluation Office (IEO) of the Global Environment Facility (GEF), focusing on adaptation taking place at the nexus between nature and humanity. I also outline a novel framework for assessing additionality, which brings together the intended impacts of GEF's work on the environment and societal dimensions.

Implications for Evaluation

All interventions take place in an environment that encompasses the natural (biophysical) and human (social, cultural and economic) spheres. The SDGs are intended to provide an integrated perspective, with all 17 giving due consideration to the three pillars of sustainable development (social, economic and environmental) (see e.g. Griggs et al. 2017), although in practice, in policymaking and in evaluation, the focus is almost exclusively on the economic benefits. The social dimension receives some attention, but mostly in terms of how it supports the economic in terms of matters such as labour productivity. The environmental is virtually ignored or receives lip service as an afterthought even though all development depends upon it (Reid et al. 2017).

Furthermore, the SDGs in practice may easily lead to new silos as organizations claim stakes in addressing particular goals. This can be seen, for instance, in the United Nations system, in which primary responsibility for the various SDGs has been carved out for specific agencies; the Food and Agriculture Organization of the United Nations focuses on Goal 2 (No hunger), whereas UN Women's mandate is with Goal 5 (Gender equality), and the United Nations Educational, Scientific and Cultural Organization's is with Goal 4 (Quality education). In practice, all of the SDGs are closely linked with each other (see e.g. Vladimirova and Le Blanc 2015). The agencies working towards achievement of their mandated SDGs often recognize

the interlinkages but may lack the resources and skills to extend their work beyond their narrowly defined mandate.

Evaluation as a practice and profession has its roots firmly in social inquiry and econometrics, both approaches focusing on attempting to measure the effectiveness of discreet interventions. Favoured approaches have included experimental and quasi-experimental designs, scenario building and cost–benefit analyses, which have been presented as the gold standard for evaluation¹. These approaches have been widely criticized for their lack of explanatory power, external validity and appropriateness and their ethical challenges, in particular in the case of international development (see e.g. Ngii 2020; Abimbola 2020). Many approaches to evaluation have been used, and significant progress has been made in inclusiveness, gender rights and human rights (see e.g. UNEG 2016), although the prevalent approach to evaluation is still narrowly focused on projects, evaluating against their internally defined logic models instead of placing them into context (Patton 2020b). By context, I mean the context in which the evaluation takes place and, more importantly, the context of the evaluand and how it relates to its societal, political, historical and cultural setting. Incorporation of biophysical dimensions has lagged seriously behind in mainstream evaluation practice. Evaluating in the coupled human–natural systems is a necessity for sustainability-ready evaluation (Rowe 2019).

On the other hand, there have long been efforts to assess the effectiveness of environmental interventions from the natural science perspective. These, for their part, have sometimes left out the human dimension. Even in the GEF, earlier evaluations tended to focus exclusively on outcomes of such things as biological diversity; greenhouse gas emissions and ecology of lakes, rivers and coasts, with little consideration for what happened to the people living in the project areas. In the GEF, a shift can be detected around the mid-2000s, which coincided with a broader realization in conservation circles that environmental protection that ignored local development interests would be doomed to failure (see e.g. GEF EO 2006). Since then, there has been a marked shift in GEF strategy towards addressing the drivers of environmental change, which can be found in the economic and development sectors.

There are two important, interlinked implications for evaluation if we as evaluators are going to contribute successfully to transformational change.

¹ As late as 2019, the prestigious Nobel Prize in economics was awarded to three proponents of experimental designs in development evaluation: Abhijit Banerjee, Esther Duflo and Michael Kremer (Royal Swedish Academy of Sciences 2019).

First, we must move beyond the project mentality in which the focus of evaluation is on measuring the effectiveness of individual interventions as if they existed in a vacuum (Feinstein 2019; Magro and Van den Berg 2019; Patton 2020b; Uitto 2016). Every intervention, whether it is a project, programme or policy, takes place in a broader landscape where it interacts with other forces and actors, some of which may be reinforcing, whereas others often work counter to the aims of the intervention. Furthermore, the relationship between the intervention and its environment may be manifold; the intervention may be targeted to change (parts of) the environment, or its goal may be outside of the immediate environment, and it may have unintended effects on the environment that may be positive or negative. In such a situation, it is futile simply to check boxes regarding whether the intervention produced the outputs it set out to produce without analysing whether it made a difference in the larger system that it is part of.

Evaluations must also seek to identify and understand unintended and unforeseen consequences that the intervention's logic model will not capture. It is safe to assume that every intervention will have environmental consequences, whether intended or unintended, positive or negative. Similarly, it will not always be possible to identify win-win solutions for all groups, which may also have different priorities and goals for an area or use of a resource (Rowe 2012). Therefore, evaluators must be able to point out possible conflicts while being sensitive to power relations and differences in vulnerabilities between groups (including indigenous peoples). Individual evaluations will not be able to produce all the necessary analytical knowledge, but it is incumbent upon evaluators to draw on scientific knowledge and to collaborate with researchers in identifying synergies and managing trade-offs (Bierbaum et al. 2018).

Because such broader contexts consist of human and natural elements, evaluations must systematically consider both systems. This nexus between nature and humanity is where sustainable development happens, if it is to happen. The need to incorporate natural and human systems, identifying synergies and trade-offs, demands a change in focus in evaluation approaches and the methods we use. Instead of evaluating against logic models of projects, we need open theories of change that place the intervention into the broader context, take into account other parts of the complex system and how the intervention interacts with them and are open to detecting unanticipated consequences. These are still theory-based approaches to evaluation but applied at a higher level.

The approach also calls for choosing evaluation approaches and methods depending on the questions to be answered. The standard

approach is a mixed-methods one that may encompass quantitative and qualitative methods. In the case of the GEF, working in environmental and natural resources management domains has allowed us to use remote sensing and geospatial methods effectively in combination with a range of other methods (Lech et al. 2018; Runfola et al. 2020; Sidman, Batra and Fuhrig 2020). An open theory of change combined with a mixed-methods approach allows us to evaluate GEF-funded projects and programmes in terms of their relevance, effectiveness, impact and sustainability; that is, do they make a difference in the global environmental problems that are the focus of the GEF's work, do they do so in a way that also benefits the people who depend on the landscape and resources where the interventions take place and are the benefits sustainable? I also ask what GEF's additionality in these situations is. Herein, I demonstrate the above points in light of practical examples from recent GEF evaluations.

Adaptation

An area squarely in the nexus between natural and human systems is adaptation to climate change. As global climate change has accelerated and national commitments under the Paris Agreement (even if fully implemented, which seems extremely unlikely) have been inadequate to halt warming below the target 2°C, adaptation has gained increasing urgency at the policy and practical levels. This by no means implies that mitigation efforts should be abandoned as futile, just that they need to be complemented by actions to adapt. The Global Commission on Adaptation calls adaptation a human, environmental and economic imperative (GCA 2019).

On the human level, adaptation solutions must address power structures and dynamics because climate change exacerbates inequality between the rich and the poor and puts a disproportionate burden on women. It tends to be the people living closest to the land – such as small farmers (a large proportion of whom are women) and Indigenous peoples – who are most vulnerable to the effects of climate change. As for the environmental imperative, degradation of the natural environment – including loss of biodiversity and ecosystem integrity – removes many of the protections that the natural environment provides against climate-related and other environmental hazards, including cyclones and storm surges, floods, droughts and heatwaves. Finally, according to the Global Commission on Adaptation, economic returns to investments in resilience are very high. Adaptation brings multiple benefits.

The costs of inaction will be dramatic, threatening cities, especially in coastal areas, from New York to Tokyo to Lagos, as well as global food security. Again, poor and vulnerable people and countries will bear the brunt of the immediate costs in terms of loss of life and livelihood.

In the field of evaluation, one of the earliest efforts to bring together the emerging community of climate evaluators around the topic of adaptation was the 2008 International Conference on Evaluating Climate Change and Development held in Alexandria, Egypt (Van den Berg and Feinstein 2009). Since then, evaluators have made progress (see e.g. Bours, McGinn and Pringle 2015), but there are still no widely accepted standards or benchmarks against which to measure adaptation. Groups such as the Technical Evaluation Reference Group of the Adaptation Fund and the Green Climate Fund Independent Evaluation Unit are working to find solutions to the evaluation challenges. One in particular relates to the nature of adaptation outcomes; successful adaptation often means the absence of something negative (e.g. a natural disaster, loss of a harvest). Evaluators would thus have to evaluate against a hypothetical counterfactual of what might have happened without the intervention.

As an operating entity of the financial mechanism of the United Nations Framework Convention on Climate Change, the GEF plays an important role in financing climate change adaptation in developing countries. The GEF and its network of partners have developed a financing framework based on the concept of climate-resilient development (GEF 2016, 21). At the United Nations Framework Convention on Climate Change Conference of Parties in Marrakech, Morocco, in 2001, three new avenues for adaptation funding were established: the Least Developed Countries Fund, the Special Climate Change Fund and the Adaptation Fund. The GEF directly manages the first two of these, whereas the Adaptation Fund is a separate entity with its own governance mechanism that the GEF administratively supports. In addition, in response to a Conference of Parties request, the GEF launched the Strategic Priority for Adaptation to pilot and demonstrate activities to reduce vulnerability and increase adaptive capacity to climate effects in GEF's focal areas (GEF 2016). The GEF's support of adaptation has been through these windows.

Adaptation in GEF programming has focused largely on least developed countries (LDCs) and other countries that are particularly vulnerable to climate change and have limited capacity to cope. Two recent evaluations focusing on LDCs (GEF IEO 2020a), and specifically on the Sahel and Sudan–Guinea Savanna Ecosystems in Africa (GEF IEO 2020b), found that adaptation to climate change featured centrally in the GEF portfolios in

these groups of countries. Thirty-four per cent of GEF funding in LDCs was allocated to adaptation, and 23 per cent of all GEF funding and 78 per cent of climate change funding went to adaptation in the two African biomes. These figures reflect the importance of the topic for these poor countries. Although their contribution to climate change has been minimal in comparison with that of more industrialized countries, they (alongside small island developing states, a few of which are also LDCs) are bearing the brunt of its impacts. GEF additionality lies in integrating adaptation to climate change into development plans and programmes.

Enhancing resilience to climate shocks is essential. Resilience can be seen as incremental (adaptive) or transformative change. The former refers to various adjustments that people or communities make to adapt to changing conditions and may include new agricultural techniques or farming practices, diversified livelihood strategies and social organization. Transformative change involves more-fundamental systemic shifts, for example, when a region changes its economic strategy. These shifts may include a combination of technological innovations, institutional reforms, behavioural shifts and cultural changes (Pelling 2011).

The interventions that the GEF has supported in the case study countries have varied considerably because adaptation is by definition place specific. What is common is the focus on ecosystem-based adaptation (using ecosystem restoration to reduce the vulnerability of human social and economic systems to climate impacts). For instance, the regional project *Adaptation to Climate and Coastal Change in West Africa – Responding to Shoreline Change and Its Human Dimensions in West Africa through Integrated Coastal Area Management*, which the United Nations Development Programme (UNDP) implemented, addressed coastal dune sustainability, which is a major environmental problem in Mauritania. The project piloted a method of reconstituting the ecosystem and biodiversity of part of the coastal dune, making it possible to secure the Mauritanian capital of Nouakchott against ocean incursion. The project *Integrating Climate Resilience into Agricultural Production for Food Security in Rural Areas*, which the Food and Agriculture Organization of the United Nations implemented in Mali, contributed to greater resilience of local grain production systems, diversification of revenue sources for rural communities, training, and restoration of soil fertility through climate-resilient techniques. Similarly, in Cambodia, the UNDP-implemented project *Promoting Climate-Resilient Water Management and Agricultural Practices* introduced new technologies, such as solar pumps, and adaptive agricultural practices that improved the livelihoods of farmers.

Overall, the two evaluations found that GEF programming has been relevant to these countries' priorities in the nexus between natural and human systems. The LDC evaluation also found that adaptation projects performed better on average than projects in other GEF focal areas.

The countries face immediate challenges pertaining to climate change and other environmental impacts that affect the lives and livelihoods of people and communities. Addressing these requires interventions in the natural environmental sphere and the social and economic spheres. The recent shift in GEF strategies towards greater integration has not decreased its relevance. On the contrary, identifying and addressing the root causes of environmental change allows the GEF to address fundamental environment and development challenges effectively.

Additionality

Since its inception in connection with the United Nations Conference on Environment and Development in Rio de Janeiro in 1992, the rationale of the GEF has been to catalyse action to generate global environmental benefits. As highlighted in the above discussion, environment and development are closely related. Because environmental protection takes place where people live and where their activities affect the natural environment, it is not possible to achieve environmental goals while ignoring people. The first study to examine this systematically in the GEF context confirmed this overall conclusion empirically (GEF EO 2006). The GEF strategy to achieve the global environmental benefits it seeks is to address the drivers of environmental degradation in human systems. Important ways involve seeking win-win solutions for people and the environment and instigating legal, policy and regulatory reforms that are beneficial for the environment.

Like other multilateral financial institutions, the GEF is concerned about whether its funding is truly incremental and not displacing other funding. The GEF's additionality was originally formulated in terms of incremental cost (the difference between business as usual and the additional cost of achieving these developmental benefits in an environmentally sound way), which the GEF would fund. This question is closely related to the true impact of the GEF and how much measurable change could reasonably be attributed to GEF funding. In other words, the question is that of a credible counterfactual: Would these changes have happened without GEF funding? Following an evaluation in 2006 that found much confusion and weak understanding of the application of the incremental cost

principle, the GEF Council simplified determination of incremental cost the following year (GEF 2007). The incremental cost analysis continued to focus on the global environmental benefits, ignoring the human aspects. The new additionality framework that the IEO developed would allow for systematic capture of the GEF's additionality through its policies, strategies, portfolios, programmes and projects (GEF IEO 2018a).

Based on a review of policies and practices of other agencies (including those of the multilateral development banks) and academic literature (e.g. Bennett 2010; Valatin 2012), the IEO framework proposed adoption of six areas of GEF's additionality: specific environmental, legal and regulatory, institutional and governance, financial, socioeconomic, innovation.

The framework would allow for better capture of the GEF's additionality across the different domains, not only the direct environmental benefits. As the fundamental justification for establishing the GEF, the global environmental benefits have been the focus of programming, as well as monitoring and evaluation systems, although there is evaluative evidence that this narrower focus has underestimated the GEF's broader impacts in the environment–development nexus (GEF IEO 2018b, 2018c). Furthermore, IEO analysis suggests that additionality beyond direct environmental outcomes is not fully understood within the partnership. For instance, the GEF most often achieves its environmental goals through engaging in legal and regulatory reform or institutional and governance improvement. Working in the coupled human–natural systems, GEF's work affects the people who depend on the ecosystem for their livelihood, be they farmers or fishermen, or whose actions affect environmental sustainability. Conversely, the drivers of environmental destruction – biodiversity loss, deforestation, land degradation, fisheries depletion, chemical pollution, climate change – are in the economic sector. To be effective, the GEF must work in the productive sectors to address the root causes. Finding win-win solutions and identifying trade-offs is necessary. Finally, since its inception, innovativeness has featured centrally in the GEF's strategies.

The IEO has applied the framework to evaluations that have been conducted since, including the evaluation of GEF support to biodiversity mainstreaming (GEF IEO 2019). Mainstreaming refers to integrating biodiversity into broad development policy, planning and practice as a mechanism to address the drivers of biodiversity loss while achieving multiple environmental and development goals. The Convention on Biological Diversity has recognized the mainstreaming goal as important, but its operationalization has been challenging (UNEP 2010). Biodiversity mainstreaming has gained in importance in GEF programming over the years,

and mainstreaming continues to be a strategic objective under the biodiversity focal area of the GEF. The current emphasis on programmatic and integrated approaches at the landscape and seascape levels reflects the importance of mainstreaming biodiversity into productive sectors, as well as in various environmental domains. Inclusion of natural capital assessment and accounting as a GEF priority is a significant step in making the business case for biodiversity. By definition, biodiversity mainstreaming takes place at the nexus of natural and human systems.

The goal of the evaluation was to assess the effectiveness of GEF contributions to biodiversity mainstreaming and to identify good practices and challenges in biodiversity mainstreaming interventions. The evaluation used mixed methods. Recognizing that country context and external variables that are outside the influence of most projects very much determine the extent of mainstreaming, the evaluation focused on three countries selected based on their representativeness of the opportunities and challenges in mainstreaming; Colombia, India and South Africa are lower- to upper-middle-income countries that have established governance frameworks and capacities for environmental management. The country studies examined biodiversity mainstreaming in productive economic sectors (mining, coffee, cattle ranching, grape cultivation, fisheries), as well as geographically in relation to land management and sustainable resource management practices.

I will not dwell on the overall findings or the country-specific results of the evaluation but instead will focus on the specific analysis of GEF's additionality. The evaluation, which was one of the first two in which we used the additionality framework, demonstrated the framework's utility and the insights it could bring in terms of, especially, the areas where human systems meet biodiversity.

The evaluation found that the GEF biodiversity mainstreaming portfolio has contributed to the various dimensions of additionality, including legal and regulatory, institutional and governance, financial, socioeconomic, and innovation, in addition to the specific environmental additionality. These include innovative approaches based on multi-stakeholder partnerships linking grassroots organizations to regional research institutions, advocacy platforms and national environmental authorities. Landscape management practices have been validated and have then influenced national policy and legislative and regulatory reform. Several projects were found to have contributed to important biodiversity legislation; transformed core institutional and sector practices and led to measurable conservation impacts in forest cover, pasture and other biodiversity indicators.

Still, systematically quantifying the social and economic benefits of biodiversity mainstreaming is difficult, although the evaluation identified plausible cases in which the projects had generated such benefits. For example, in Colombia, coffee growers who adopted environmentally sound shade cultivation and agroforestry practices received better prices for their produce, which in turn resulted in higher incomes. Similarly, in India, the Sustainable Land Management in Shifting Cultivation Areas of Nagaland for Ecological and Livelihood Security project that the UNDP implemented benefited more than 3,000 women, whose income from sales of produce from the *jhum* cultivation system rose 25 per cent during the project period. In the project area, 78 per cent of surveyed farmers felt that their agricultural income had increased during the project period. Systematically quantifying such socioeconomic benefits will be a future priority.

Conclusions

The close interdependence between the natural and human systems is recognized more widely than perhaps ever before. The COVID-19 pandemic has underscored that humans are not separate from the Earth's ecological system. The way we exploit and abuse the natural environment contributes significantly to the increasing occurrence of zoonotic pandemics as human activities encroach deeper into natural ecosystems and we come into closer contact with non-human animals. Evaluation can play an important role in uncovering evidence from past experiences and demonstrating the importance of maintaining ecosystem integrity and a stable climate, not only for the purposes of the environment, but also for human welfare and health (GEF IEO 2020c).

Still, as a community and a profession, evaluation is not yet in a place where it can effectively address sustainability. Evaluators struggle with coupled human and natural systems. We are still stuck in a project-centred mindset in which we tend to evaluate interventions in isolation against their internal logic, although there are clear signs that the situation is changing. The discussion about the need for systems approaches to evaluation has found its way to many conferences, listservs and communities of practice where evaluation is discussed. Many evaluators and evaluation users recognize the need, but the practice lags. The intervention focus dominates most of evaluation practice, largely driven by funders' demand for accountability. In a complex system in which attribution of specific changes to an individual intervention is difficult, such accountability focus can be counterproductive

and could be achieved through tools other than evaluation (e.g. performance audit). It is important for evaluators to place the evaluand into the context in which it operates and, specifically, in which it interacts with human and natural systems. This will require an open theory of change that pays attention to unanticipated consequences – to the environment; to different groups of people, especially the most vulnerable; to incentives and disincentives for sustainability – and whether the intervention makes a positive difference in the problem it was established to address.

Most development takes place at the nexus between nature and humanity. This relationship has been heightened as we have entered the Anthropocene, an era in which human impacts are so pervasive that they lead to significant modifications in the Earth's biosphere and geosphere. Although we seek win-win solutions in which people and the planet both benefit, these are not always easy to find. Evaluators must be clear on the choices and trade-offs that may be necessary. It is imperative for evaluators to remain relevant to rise to the challenge of evaluating as if both people and the environment mattered.

Acknowledgements

IEO staff Anupam Anand, Geeta Batra, Carlo Carugi and Anna Viggh managed the GEF evaluations referred to in this chapter.

References

- Abimbola, S. 2020. Evidence as cliché: Using trials to tamper with governance. *CODESRIA Bulletin* 4: 9–15.
- Asokan, G.V., and V. Asokan. 2016. Bradford Hill's criteria, emerging zoonoses, and One Health. *Journal of Epidemiology and Global Health* 6: 125–129.
- Bennett, K. 2010. Additionality: The next step for ecosystem service markets. *Duke Environmental Law and Policy Forum* 20: 417–438
- Bierbaum, R., A.L. Cowie, R.O. Barra et al. 2018. Integration: To solve complex environmental problems. Washington, DC: Scientific and Technical Advisory Panel to the Global Environment Facility. <https://tinyurl.com/w9ddm9xe>
- Bours, D., C. McGinn and P. Pringle, eds. 2015. Monitoring and evaluation of climate change adaptation: A review of the landscape. *New Directions for Evaluation* 2015 (145).
- Buchner, B., A. Clark, A. Falconer et al. 2019. Global landscape of climate finance 2019. Climate Policy Initiative: London. <https://tinyurl.com/4hpvzx3>

- Burke, M., S. M. Hsiang and E. Miguel. 2015. Climate and conflict. *Annual Review of Economics* 7: 577–617.
- Caballos, G., P.R. Ehrlich and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*. www.pnas.org/cgi/doi/10.1073/pnas.1704949114
- Feinstein, O. 2019. Dynamic evaluation for transformational change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R.D. van den Berg, C. Magro and S. Salinas Mulder, 17–31. Exeter, UK: IDEAS.
- Garcia, J. R., and A. Zazueta. 2015. Going beyond mixed methods to mixed approaches: A systems perspective to asking the right questions. *IDS Bulletin* 46 (1): 30–43.
- GCA (Global Center on Adaptation). 2019. Adapt now: A global call for leadership in climate resilience. Rotterdam and Washington, DC: Global Commission on Adaptation and World Resources Institute. <https://tinyurl.com/dybkdxvp>
- GCA (Global Center on Adaptation). 2021. The Global Commission on Adaptation. <https://gca.org/global-commission-on-adaptation/home>
- GEF (Global Environment Facility). 2007. Operational guidelines for the application of the incremental cost principle. <https://tinyurl.com/5cb7w2ek>
- GEF (Global Environment Facility). 2016. Time to adapt: Insights from the GEF's experience in adaptation to climate change. Washington, DC: GEF. <https://tinyurl.com/22pd2mzm>
- GEF EO (Global Environment Facility Evaluation Office). 2006. *The role of local benefits in global environmental programs*. Washington, DC: GEF EO. <https://www.gefio.org/evaluations/lb-global-environmental-programs-2006>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2018a. An evaluative approach to assessing GEF's additionality. <https://www.gefio.org/evaluations/additionality-framework>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2018b. Evaluation of the multiple benefits of GEF support through its multifocal area portfolio. Washington, DC: GEF IEO. <https://www.gefio.org/evaluations/multiple-benefits-2016>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2018c. *Impact of GEF support on national environmental laws and policies in selected countries*. Washington, DC: GEF IEO. <https://www.gefio.org/evaluations/regulatory-reform-2017>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2019. *Evaluation of GEF support to mainstreaming biodiversity*. Washington, DC: GEF IEO. <https://tinyurl.com/ywmcnufe>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2020a. *Strategic country cluster evaluation of the least developed countries*. Washington, DC: GEF IEO. <https://www.gefio.org/evaluations/scce-ldc>

- GEF IEO (Global Environment Facility Independent Evaluation Office). 2020b. *Strategic country cluster evaluation: Sahel and Sudan–Guinea savanna ecosystems*. Washington, DC: GEF IEO. <https://www.gefio.org/evaluations/scce-biomes>
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2020c. *The GEF response to crisis: What can we learn from evaluations?* Washington, DC: GEF IEO. <https://www.gefio.org/sites/default/files/documents/ieo-covid-2020.pdf>
- Griggs, D.J., M. Nilsson, A. Stevance et al., eds. 2017. *A guide to SDG interaction: From science to implementation*. Paris: International Council for Science. <https://tinyurl.com/3h494t5d>
- IPCC (Intergovernmental Panel on Climate Change). 2018. Special report: Global warming of 1.5°C. <https://www.ipcc.ch/sr15/>
- Lech, M., J. I. Uitto, S. Harten et al. 2018. Improving international development evaluation through geospatial data and analysis. *International Journal of Geospatial and Environmental Research* 5 (2). <https://dc.uwm.edu/ijger/vol5/iss2/3/>.
- Magro, C., and R. D. van den Berg, 2019. Systems evaluations for transformational change: Challenges and opportunities. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R. D. van den Berg, C. Magro and S. Salinas Mulder, 131–155. Exeter, UK: IDEAS.
- Ngii, D. 2020. The *Randomista* rampage in development economics: A view from the trenches. *CODESRIA Bulletin* 4: 3–7.
- Patton, M.Q. 2020a. Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation* 1–37. <https://doi.org/10.1177/1098214020933689>
- Patton, M.Q. 2020b. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Pelling, M. 2011. *Adaptation to climate change*. New York: Routledge.
- Reid, A. J., J. L. Brooks, L. Dolgova et al. 2017. Post-2015 Sustainable Development Goals still neglecting their environmental roots in the Anthropocene. *Environmental Science and Policy* 77: 179–184.
- Ripple, W.J., and R. L. Beschta, 2012. Trophic cascades in Yellowstone: The first 15 years after wolf reintroduction. *Biological Conservation* 145 (1): 205–213.
- Rowe, A. 2012. Evaluation of natural resource interventions. *American Journal of Evaluation* 33 (3): 384–394.
- Rowe, A. 2019. Sustainability-ready evaluation: A call to action. In: *Evaluating sustainability: Evaluative support for managing processes in the public interest*. *New directions for evaluation*, (ed.) G. Julnes, 162: 29–48. San Francisco: Jossey-Bass.
- Royal Swedish Academy of Sciences. 2019. Press release: The prize in economic sciences 2019. <https://tinyurl.com/2a3cn252>

- Runfola, D., G. Batra, A. Anand et al. 2020. Exploring the socioeconomic co-benefits of global environment facility projects in Uganda using a quasi-experimental geospatial interpolation (QGI) approach. *Sustainability*. <https://www.mdpi.com/2071-1050/12/8/3225>
- Sidman, G., G. Batra and S. Fuhrig. 2020. The use of remote sensing analysis for evaluating the impact of development projects in the Yellow Sea large marine ecosystem. *Sustainability*. <https://www.mdpi.com/2071-1050/12/9/3628>
- Uitto, J.I. 2016. The environment-poverty nexus in evaluation: Implications for the Sustainable Development Goals. *Global Policy* 7 (3): 441–447.
- UN DESA (United Nations Department of Economic and Social Affairs). n.d. Transforming our world: The 2030 Agenda for Sustainable Development. <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- UNEG (United Nations Evaluation Group). 2016. Norms and standards for evaluation (2016). <http://www.unevaluation.org/document/detail/1914>
- UNEP (United Nations Environment Programme). 2010. Decision adopted by the Conference of the Parties to the Convention on Biological Diversity at its tenth meeting. <https://www.cbd.int/doc/decisions/cop-10/cop-10-dec-02-en.pdf>
- UNEP (United Nations Environment Programme). 2016. Frontier 2016: Emerging issues of environmental concern. Nairobi: UNEP. <https://tinyurl.com/c98afr4b>
- UNFCCC (United Nations Framework Convention on Climate Change). 2021. The Paris Agreement. <https://tinyurl.com/3twkkcnd>
- United Nations. 2018. Secretary-General's remarks on climate change [as delivered]. <https://tinyurl.com/44kcv69k>
- Valatin, G. 2012. Additionality and climate change mitigation by the UK forest sector. *Forestry* 85 (4): 445–462. <https://academic.oup.com/forestry/article/85/4/445/519175>
- Van den Berg, R. D., and L. 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.) J.I. Uitto, J. Puri and R. D. van den Berg, 13–34. Dordrecht: Springer. https://link.springer.com/chapter/10.1007/978-3-319-43702-6_2
- Van den Berg, R. D., and O. Feinstein (eds.). 2009. *Evaluating climate change and development*. London and New York: Routledge.
- Vladimirova, K., and D. Le Blanc. 2015. How well are the links between education and other Sustainable Development Goals covered in UN flagship reports? A contribution to the study of the science-policy interface on education in the UN system. DESA Working Paper No. 146. UN DESA. https://www.un.org/esa/desa/papers/2015/wp146_2015.pdf



CHAPTER 6

Contradictions and Complementarities Between South and North on Transformation in the Anthropocene

MEHJABEEN ABIDI-HABIB, JANE BURT, JOHN COLVIN,
CHIMWEMWE MSUKWA AND MUTIZWA MUKUTE

Abstract. Emerald Network is an emerging community of evaluation and learning praxis working in the field of international cooperation and development and with Global South consultancy partners. Our evaluation and learning praxis draws on our combined experience in policymaking, design, strategy, finance, implementation and research. Recognizing that we are living through the early Anthropocene – or Capitalocene to be more precise – we seek to contribute to transformative development pathways in service to a just, regenerative, low-carbon, resilient world. In this chapter, we reflect on how our praxis has evolved over the past eight years, sharing stories of success and failure and what we have learned in the service of transformational work. The purpose of this chapter is to explore the role of evaluation praxis in transformational design for sustainable development, focusing on a number of themes that have come to play a central role in our praxis. These include navigating and learning through contradictions and complementarities between Global South and Global North, the centrality of navigating power in these contradictions and complementarities, the value of understanding history and context, the importance of internal praxis and the design and facilitation of adaptive and potentially transformational learning processes.

Introduction

Much has been written about the need for transformation in the Anthropocene, seeking to define transformation in this context – a context that should be better known as the Capitalocene. Having worked to support the Climate Investment Funds (2019) in developing an understanding of transformation practice, we draw on this framework, which embraces four dimensions, all of which the Climate Investment Funds see as necessary to define transformation:

- *Transformation has relevance.* For Emerald Network, the relevance of our transformational praxis¹ is to contribute to social justice, ecological flourishing and well-being for all.
- *Transformation involves systemic change.* Transformation involves fundamental changes in structures and systems – disrupting these systems and unlocking new pathways to development – and is facilitated through systemic practices.
- *Transformation takes place at scale.* Transformation involves working at scale and taking to scale. Strategies include multiscale development, scaling up, scaling out and scaling deep.
- *Transformation is an enduring, sustainable process.* Transformation involves a long-term, dynamic process that builds capacity to ride out short-term shocks and transcend longer-term stresses.

Transformation as Systemic Change

In this chapter, our primary focus is on the second dimension of transformation as systemic change and, along with this, the value of leveraging systemic change through systemic interventions. (Alongside this primary focus, the other three dimensions of transformation remain integral to our praxis, and we touch on these in various ways throughout the chapter.)

In foregrounding systemic change, this chapter explores five interlinked themes (complementarities and contradictions, centring power, context and history, designing for transformational learning, internal and external design praxis), which we engage with through the following four lenses. First, we briefly introduce each theme. Second, theoretical background

¹ We refer throughout to our evaluation and learning practice as ‘praxis’, in the sense of theory-informed practice. We understand evaluation and learning to be shaped by theory as well as experience and work reflexively with this understanding.

on each theme is provided in boxes 6.1 to 6.5. Third, illustrations of these themes as they illuminate particular stories of Emerald Network praxis can be found in the sections that follow, in the latter, Joint Reflections from Our Community of Praxis part of each story. Fourth, in the final section on concluding insights, we return to each of these themes in a discussion of transformational design for evaluation. The reader is invited to review these themes through each of these four lenses.

We start with a brief introduction of each theme.

- *Complementarities and contradictions.* Systems comprise multiple elements, often with complex relationships and feedback loops between them. Working to change system structure and function involves working to shift these relationships, sometimes in radical ways. One approach to thinking about inter-relationships is to work with complementarities and contradictions between the elements – a central theme of this chapter. Complementarities and contradictions are inherent not only in our systems of interest, but also in our own collaborations, where their consideration is integral to what we refer to as internal design praxis (in the fifth theme below).
- *Centring power.* Working with complementarities and contradictions between elements of a system immediately takes us into issues of power. Centring and consciously facilitating power and political relationships and factoring these into transformational design praxis is a second core theme of this chapter and informs our reflections on relationships between the Global South and the Global North – as geographical spaces and as metaphors for unhealthy power disparities.
- *Context and history.* Working with complex systems and their inherent power dynamics requires paying careful attention to context and history – recognizing especially that solutions for complex systems are always specific to their context and history, much more so than for complicated or simple solutions.
- *Designing for transformational learning.* Recognition of the dynamics of complex systems, and the many uncertainties and emergent properties inherent in these, calls for design and evaluation approaches that foreground adaptive learning. In the case of evaluation for transformation – and transformative evaluations – this may also call for designing transformational learning processes.

Box 6.1 Theorizing Complementarities and Contradictions

Complementarities. The concept of solidarity, which is characterized by relatedness or *ukama*, an essential component of the African philosophy of *ubuntu*, underpins our work with the concept of complementarities. *Ukama* encourages being, thinking and doing with others in the service of the common good and linking past, current and future generations in creating and sharing knowledge, moral values and natural heritage (Murove 2009). We work with three kinds of solidarity that are essential for building complementarities between the Global North and Global South: relational solidarity, which is committed to reciprocity and the act of being with others as part of them; transitive solidarity, which involves taking action to change the way things are in a reflexive process that transforms the agent in the processing of acting and creative solidarity, which is collective learning to reveal new horizons and produce new ways of being together, of making, feeling, creating and loving (Gaztambide-Fernandez 2012).

Contradictions. Engeström (2001) theorizes contradictions as historically accumulated structural tensions within and between practices, actors and groups of actors that are available as important sources of transformative learning, change and development. Contradictions are by nature tension laden. They are subterranean and invisible, have a history, are structural and require surfacing, for example from relationships of political and institutionalized power, choice and decision-making power at different scales, and access to and control of material and non-material resources such as knowledge. Confronting contradictions is uncomfortable, but failure to resolve them can worsen situations, so confronting them is critical for bringing about transformational change that goes beyond addressing symptoms.

Box 6.2 Theorizing Power

The feminist concept of power, which emphasizes *power with* and *power to* as opposed to *power over* others, has sensitized our evaluation work and our relationships as co-evaluators (Karlberg 2005). This feminist model of power is against conflating power with domination but views power as the human ability to act in concert with others, nurturing and empowering others to produce change (Arendt 1969). We also draw on system theorists such as Boulding (1990) who encourage working with integrative power, which is underpinned by cooperation and reciprocity, a sense of community and the ability to create and pursue desirable things together.

Box 6.3 Theorizing History and Context as Part of Transformative Praxis

As our internal praxis matures, the team's experiences converge around common value-based threads. A core perspective for transformation praxis is *centring on the local*, which we used to navigate our way into the Swedish Expert Group for Aid Studies (EBA) assignment (below), a work of large, multiscalar scope. In centring the local, the impact of change-making becomes one strand within place-specific context and history: a strand that surfaces power legacies, complexity, surprise and fine-tuned responsiveness from outside drivers such as donors and intermediaries.

Theory informs us 'that choices between possible pathways, at different scales and for different groups of people, are shaped by uneven power structures and historical legacies that create their own, often unforeseen change...and that considerations for inclusiveness, place-specific trade-off deliberations, redistributive measures and procedural justice mechanisms [must] facilitate equitable transformation' (Roy et al. 2018).

Box 6.4 Theorizing Design for Transformational Learning

That ongoing cycles of learning and adaptive management are co-designed, are guided by essential principles and offer a tensile holding framework are some of the emerging insights from our journey of eight years of collaborations towards transformational design.

Theory casts light on the imperative for 'creating environments that enable learning and knowledge management...[with] learning increasingly understood to be the linchpin [of good evaluation]' (STAP 2017). Furthermore, 'jointly practicing the essentials will create a highly adaptive, reflexive, relational, collaborative and impact-oriented form of [design] that has a strong impetus [for evaluation] to engage with action' (Fazey et al. 2018).

Box 6.5 Theorizing Internal and External Praxis

With our focus on transformative praxis, one of our key assumptions has been that the internal reflective praxis of the consultant team also shapes the outer system of interest. For example, in the Pakistan story below, we purposively bridged from being good consultants to facilitators of transformation while experiencing between us the power play of whose knowledge counts in the contradictions between Indigenous and state and donor priorities.

Theory informs us that 'reflective practice should not shy away from dealing overtly and reflectively with conflicts of views, values, and rationality' and that 'a greater use of reflective practice is advocated in reference not only to [internal praxis] development, but as a means to enhance dialogue, stakeholders' involvement and organisational learning' as external praxis (Kubera 2019).

- *Internal and external design praxis.* Working with transformational learning calls for paying careful attention to what we call internal praxis – processes of individual and team reflection, reflexivity, learning and development – alongside what we do in the external world. In our experience, attending to internal praxis not only enables us to be better practitioners, but can also provide us with vital clues as to what is going on in the systems we are participating and intervening in.

The five themes set out above can be considered principles of design as well as analytical themes. A principles-based approach to evaluation design is especially important in a transformational context, in turn requiring a toolkit of frameworks and approaches that we can apply, for example, to help us understand history, context, outcomes, insights and design and facilitate processes of change and learning. A number of these frameworks and approaches are discussed in this chapter – for example, theory of change, contribution analysis and learning history.

Introducing Our Internal Values and Praxis

Within Emerald Network, we see ourselves as a community of praxis that seeks to contribute to transformative learning for sustainability and social justice. In developing interventions, we deliberately draw on the distributed knowledge and experience of our team, which is multicultural and transdisciplinary. We build our praxis by learning from different streams of evaluation thinking. As a result, we are continuously working on and expanding the boundaries of our collective practice. Our ambition is to feed back into the broader evaluation system.

Structure of This Chapter

This chapter draws on a panel discussion that we designed and held for the IDEAS 2019 Global Assembly in Prague. In transcribing this performance for a book chapter, we sought to remain faithful to the narrative form and structure of the performance while adding structural clarity to increase the accessibility of interweaving descriptive, reflective and theoretical narrative, all of which combine in praxis. To increase this accessibility, we also draw on a narrative format derived from learning history (Bradbury, Roth and Gearty 2015), which is designed to enable the reader to juxtapose multiple voices – of different actors and of descriptive, reflective and theoretical

perspectives. Learning history had already informed our performance; here we take it a step further in shaping the layout of complementary written narratives.

After this introductory section, the chapter follows four stories taken from our collaborative praxis, using these stories to describe the type of evaluative work we do and to illustrate our thesis through the use of evaluator reflections. We set out descriptive narratives first, followed by reflective insights, with the main themes highlighted in italics. Summarizing these reflective insights, the main themes revealed in our transformative evaluation praxis and its role in transformational design for sustainable development are brought together in the final section of the chapter.

An Early Collaboration in Pakistan: Facilitating Power and an Unconventional Portfolio

This first story, on which Mehjabeen and John worked together for the first time, is not an evaluation story but is about our role as transformational designers for a portfolio of proposals². In this story, we were challenged to work with and seek to resolve important contradictions that arose from application of our design principles.

Researchers' Story

The context for this story is that, in 2012, the government of Pakistan approved its first climate change policy – five years in the making. Now Pakistan's Ministry of Climate Change (MoCC) wanted a portfolio of fast-track action proposals in the climate-compatible development space³ that it could take to potential donors. The MoCC therefore agreed with the regional Asia office of the Climate and Development Knowledge Network that it would commission consultancy work to develop this portfolio – work that John and Mehjabeen won as part of a North–South consortium⁴.

For our praxis, this was a formative story. In addition to working together for the first time, we had to make choices that took us across a transformational bridge and made prominent for our community the centrality of an internal praxis and how to articulate a growing understanding of how our

² For the full story, see Colvin and Abidi-Habib (2013).

³ The space of intersections between adaptation, mitigation and development.

⁴ Global Climate Adaptation Partnership led this consortium in the United Kingdom, working with Hagler Bailly in Pakistan.

internal and external work shapes the choices that we make. Our design experience and the learning that flowed from this story is set out below.

Our design drew on a set of principles around equity, ecological justice and sustainability and brought together an unlikely alliance of institutions that would not normally have collaborated. This included institutions from different sectors, for example those working toward food security, women's empowerment and forestry. It also intentionally convened individuals who were identified as leaders, innovators or climate change experts (and sometimes more than one of these). Together this alliance developed a diverse, unconventional portfolio of climate-related projects based on the long, deep traditions of their work.

The portfolio surprised us in that it opened up a power dynamic with the part of the client system that lay within the federal government and created ambivalence or what we call 'wobbles' in our whole remit. An example of questions that surfaced was: Which knowledge counts: the Indigenous learning that our unlikely alliance brought to the table or the calls that international donors made and the expert knowledge expected to support these? Also, the variation in scale of project size was puzzling to our client.

So, silencing some of the diversity in the portfolio that was presented, the client shifted what they were asking for and requested an expert-led set of tried and tested proposals. These contradictions between our design principles and client preconceptions led to an eventual collapse of our assignment, making us understand that the institutional holding framework that our client system offered was not strong enough to follow through with a transformation agenda. The MoCC was weak in intent and purpose and lacked foresight.

We had other offers to hold together the alliance we had built, as well as generative energies from the members of the alliance, but could not pursue this particular effort. Instead, we gained a reputation that led us to several more allied assignments in Pakistan.

Our process of learning to work with internal design praxis during this assignment led Mehjabeen and John to a series of observations:

- As we faced these challenges and wobbles, we had to attend to how we were working together as a team, as well as to the personal and professional challenges this work brought up – what we came to refer to as our 'internal design praxis' – for a team and within our individual practices.
- We found ourselves needing to structure our collaborative teamwork as a series of rapid adaptive management and learning cycles,

which we also shared with the client; we had to learn to work together in ways that were agile and dexterous.

- Ultimately for this assignment, because the client chose expert knowledge over facilitation of Indigenous knowing, we had to take some tough choices about who we were as professionals in this context and what the right path forward was for us.

Further insights about internal design praxis flowed from this experience. Thus, we articulated this principled choice to ourselves as a choice between being good consultants and stepping into being facilitators of design for transformation. We noted that this choice mirrored the choice that the whole system had to make to reach a transformative space. In other words, although part of the system embodied by unlikely alliances of Indigenous knowledge was ready to design for transformation, the part embodied by the MoCC and its advisors, when faced with this choice, opted to carry on with business as usual. This insight in turn sensitized us to the value of mirroring⁵ as a technique to enable us to reflect on the relationship between internal and external praxis.

Risks of losing professional reputation accompanied the high levels of complexity and uncertainty in this work, yet they demanded a transformational approach. In turn, making this choice required that we cross thresholds of fear. It was at times quite scary for us, requiring that we step into a particular kind of leadership that involved foregrounding our knowledge less, facilitating others' knowledge and creating spaces for different knowledge systems to interact.

This also sensitized us to the importance of matching the challenge of facilitating transformational work with the tensile strength of our client system and the degree of holding that this could create⁶.

Joint Reflections from Our Community of Praxis

In this first story, we note how *contradictions* are surfaced early on, as the principled portfolio rubs up against power dynamics in the federal government, leading to an awkward question: Whose knowledge counts? Here, a contradiction emerges between the indigenous knowledge that the team's

⁵ This technique is also known as shadow consulting (see Hawkins 1993).

⁶ Ronald Heifetz and Donald Laurie (1997) write about the importance of creating adequate holding frameworks in leading organizations through complex processes of change.

principled approach had invited into the system and the 'expert knowledge' that was the expectation of the MoCC and Climate and Development Knowledge Network. The team was not purposefully excluding expert knowledge but rather holding it in reserve until such time as it might usefully complement indigenous sense-making, but the contradiction between an unconventional portfolio and a conventional institutional holding framework, with the attendant *power dynamics* at play, proved to be more than the system could hold. It is also tempting to ask whether the contradiction between indigenous knowledge (an unconventional portfolio) and expert knowledge (conventional institutional practices) could be framed as a South–North contradiction.

A second theme, relevant to *transformational design and implementation*, also emerged here. It soon became clear to the team, particularly as they worked through the process of attending to contradictions, that design would need to be expressed through ongoing cycles of adaptive management and learning.

This story also highlighted the team's attendance to *internal praxis*, which was vital to their ability to adapt and called on the application of new skills, such as attending to processes of mirroring. Internal praxis was also called on to guide the team in learning new ways of working together and encountering new areas of professional and personal contradictions and constraints, for example in balancing roles as good consultants and facilitators of transformation. Again, a contradiction between Global South and Global North practices might have been at play here, even as the team sought to bridge and integrate these.

In summary, this assignment created a significant *early developmental moment* for our incipient community of praxis, foreshadowing several of the themes that we explore further in the remainder of this chapter: The team was tested and had to make difficult choices, face the risks and feel the fear – all of which were demanding of internal praxis and led to rapid adaptive management and design. The team was challenged to work with and seek to resolve important contradictions that arose from application of its design principles, which in hindsight appeared to mirror Global South–Global North contradictions in knowledge and institutional practices, as well as in professional identities.

Evaluating the Africa Climate Change Resilience Alliance: Centring Transformation and Power in Design

Our second story takes us to Africa and the evaluation of the Africa Climate Change Resilience Alliance (ACCRA). Picking up a theme from the first story, in this work, we begin to learn about combining our role of good evaluation consultants with that of taking on the responsibilities of facilitating transformation.

Evaluators' Story

This story is of an evaluation that Mutizwa and John conducted.

ACCRA was a seven-year programme with a focus on adaptation to climate change funded by the U.K. government, coordinated by a consortium of international non-governmental organizations⁷ and implemented in Ethiopia, Mozambique and Uganda (Levine, Ludi and Jones 2011).

ACCRA had three objectives:

- Build local adaptive capacity to combat climate change, understood broadly as including assets but also soft factors such as knowledge systems⁸, innovation and flexible, forward-looking decision-making
- Transform multilevel governance systems to enable, rather than constrain, development of local adaptive capacity
- Transform gender relations as part of the above to develop effective, just, gender-just local adaptation practice

The evaluation required us to look back over the seven-year history of ACCRA and to assess outcomes and institutional arrangements at national and international levels. In Ethiopia, we were also asked to look in more depth at the contribution of ACCRA to Ethiopia's transformational pathway to climate-resilient green growth.

The ACCRA evaluation had a strong learning interest. We co-designed the evaluation using a basket of evaluation approaches, summarized in table 6.1, to ensure that we covered transformational aspects (Mukute,

⁷ Consortium members were Oxfam GB, Overseas Development Institute, Care International, Save the Children Alliance and World Vision International.

⁸ For example, local climate information systems.

Table 6.1 Design Approaches Used in Evaluating the Africa Climate Change Resilience Alliance

Aspect of evaluation covered	Evaluation design approaches used: Systems thinking based					Unit of analysis
	Learning history	Expansive learning (contradictions, agency, learning)	Theory of change	Feminist concept of power	Contribution analysis and process tracing	
History and context						Interacting, nested systems at the local, national, international, local to national, and national to international levels
Voices, power and power relations						
Accountability, learning and influence						
Transformative change mechanisms and processes						
Social, climate and gender justice; adaptive capacity and governance system transformation outcomes						

Note: For each aspect of the evaluation, the shaded cells indicate the approaches used.

Colvin and Baloi 2017). We worked closely with ACCRA partners, including Oxfam GB as the evaluation lead partner, from the outset. Programme implementation was coordinated from Kampala, and the rest of programme-level coordination was done from London.

Our design insights are as follows:

- We found that the use of complementary evaluation methods helped us develop a fuller understanding of what transpired, why and to what effect.
- We established that surfacing and articulating uneven and unhealthy power relations around institutional governance

arrangements is important for transformational designs. Our client embraced these articulations and decided to take this insight forward. Although the phase 3 bid for ACCRA was unsuccessful, the client's learning informed future Oxfam policy and design thinking on resilience, including transformational aspects (Oxfam 2016).

- We also learned that, when evaluating complex systems, power relations, silences and articulations are found not only in obvious places, such as between the Global South and the Global North (e.g. Kampala and London), but also in less obvious parts of the system, such as within households and between local government and the community.

We learned that these less obvious power relations, silences and articulations appear as metaphorical aspects of Global North and Global South.

Further reflections of Mutizwa and John on internal praxis in the ACCRA work are as follows:

- *Mutizwa*: 'This was the first time that John and I had collaborated. We found the process very generative, drawing on a complementary set of practices and approaches that enabled us to produce sufficiently rigorous evidence for accountability purposes while also attending to inclusiveness of different actors and facilitation of learning events and processes.'
- *John*: 'Given the complexity of ACCRA and the multiple levels of governance to be engaged, evaluated and facilitated, inevitably there were moments of tension in our working relationship, particularly arising from uncertainties in the evaluation and/or learning process and how best to address these. At times, one or other of us might feel quite vulnerable in making particular methodological or design choices. For example, at the start of our second period of field work in Addis Ababa, I was very uncertain about how to approach case study selection and sought Mutizwa's emotional as well as intellectual support in making this selection. In facing these dilemmas, we discovered that we could draw on feminist concepts of power to guide us.'

Joint Reflections from Our Community of Praxis

Transformational change. Having the evaluation centred around transformation created an opportunity to track transformation from initial design through implementation and toward impact, including transformative change mechanisms and processes; mechanisms of social, climate and gender justice; and adaptive capacity and governance system transformation outcomes. The ACCRA programme's focus on transforming governance systems, pathways of change and community adaptive capacities largely inspired the transformational interest of the evaluation.

In terms of *internal and external design praxis*, the ACCRA story highlights a move for us into the evaluation space and into greater complexity, involving many layers of governance and raising attendant uncertainties and even vulnerabilities in praxis. Navigating these spaces effectively required that we adopt *complementary practices and approaches* to discover rigorous evidence, inclusiveness of various actors, facilitation of learning and tracking of transformation. We also benefited from the use of the Global North and Global South as literal and metaphorical dimensions of transformation internally and externally to manage the complex design framework co-constructed for this work. Between evaluation design and implementation of the design, the ACCRA story surfaces the centrality of *complementarities and contradictions* as important learning for transformation. The team pooled their practices and sought to combine them creatively in this complex evaluation but also encountered tension and contradictions between inclusivity, rigour and feasibility. Demands for accountability, learning and influencing also generated contradictions.

The ACCRA evaluation also centred *power and power relations*, in part through analysis of power relations and their transformation and in part through giving space to multiple voices and perspectives, drawing on learning history approaches – two approaches used as the team sought to work overtly with power in the evaluation design. The ACCRA story evaluation work was more explicit than our Pakistani story in theorizing power, drawing on feminist theories of power to surface voices of power and power relations.

Feminist theories of power also proved important in beginning to explore internal praxis within this new team. Already experienced individually, as the new team grappled with the challenge of sharing power in a collaboration, the feminist concept of *power with* proved critical for navigating moments of vulnerability in the face of complexity and key design decisions and for helping to articulate and begin to work through contradictions in internal praxis. This in turn enabled us personally to learn more

about what it means to address and work into power relations between the Global South and the Global North.

Malawi: Sharpening Approaches in a Different Context and with New Team Members

For our next story, we remain in Africa, focusing this time on an evaluation set in Malawi. The story again involves Mutizwa and John working with an expanded team and joined by Jane from South Africa and Chimwemwe from Malawi (Mukute et al. 2021).

Evaluators' Story

Malawi's governance systems have failed to hold government to account, causing significant leakage of public funds. Oxfam GB engaged Emerald Network to evaluate a short-term project that piloted building communities' capacities for tracking government development expenditure at the local council level. Seeking to address the failure of local governance to be accountable, a national trust financed the project, and Oxfam in Malawi and two local partner organizations jointly implemented it. The project goal was to enable three district councils to become more accountable, responsive and inclusive in managing local development resources.

The purpose of the evaluation was for Oxfam GB and Oxfam in Malawi to develop a deeper understanding of impact generated through the project's broader contribution and so be able to support programmatic learning. During the inception process, we decided to evaluate the contribution to two focal outcomes:

- District councils are accountable, responsive, inclusive and effective in managing funds.
- Women in these districts are meaningfully empowered to participate in decision-making structures and processes.

We investigated the processes by which these outcomes were generated and assessed the significance of Oxfam's contribution to them. Two learning moments that moved us closer to transformational approaches are discussed below.

In our *first learning moment*, we drew on the ACCRA evaluation to deepen the use of complementary evaluation methods. With reference to

complementarity, the client prescribed a process tracing approach that aligned with the counter-factual approaches used in the ACCRA (Ethiopia) evaluation. The team carefully adapted this approach in the context of the evaluation questions and Malawi, expanding the system of analysis to trace contribution at the national, district and community levels.

Using theory of change to make sense of the landscape of the project, we traced evidence of the outcomes we were focusing on at multiple scales of governance. Although we engaged with stakeholders at multiple scales, we centred the experience of beneficiaries, particularly women, and triangulated this with the voices of government officials – not the other way around. We listened to the beneficiaries, particularly women, to understand how they were experiencing the effectiveness of the governance systems and the project in their own lives. We also adopted a utilization focus to bring clarity to whom the evaluation was for. It was agreed that the evaluation was for Oxfam in Malawi and its local partners. The intention was for insights emerging from the evaluation to be of value, first and foremost, to Oxfam in Malawi and local partners, even though the client was Oxfam GB.

Enabling this complex design, some team members were more experienced with the process tracing approach than others. We set up a dialogue platform within the team that enabled crossover and sharing of experience even within the short time frame during which team members met in Malawi and shared roles and responsibilities based on our individual capabilities. For example, Mutizwa examined national capacity and shared what he learned with Chimwemwe and Jane before we undertook fieldwork in a district. Chimwemwe and Jane knew their roles without having to name them. Chimwemwe, with his deep knowledge of the context and networked relationships with communities, led the fieldwork, and Jane listened for patterns with an eye on gender.

Some contradictions we encountered were that, although our initial intention was for the evaluation insights to catalyse learning for Oxfam in Malawi and local partners, the requirements of the contract quickly led to most dialogues taking place with Oxfam GB. Country-based organizations viewed this as an Oxfam GB evaluation that they were saddled with, which led to the Global North partner being more interested in the results and learning than country-based partners were. The Emerald Network evaluation team was unable to shift this particular power dynamic, which was embedded in the broader Oxfam system and was related to who has the power to call for evaluations, including when they take place. Although Oxfam in Malawi was involved in formulating the two focal outcomes, the

evaluation formed part of Oxfam GB's larger organizational undertaking to better capture and communicate the effectiveness of its work.

In our *second learning moment*, we worked with feminist concepts of power. With reference to *complementarity*, we found that working with these feminist concepts within the process tracing approach could be deepened in the Malawi case. We explored the contribution the project had made to women's empowerment, in line with the second focal outcome. Our feminist positioning enabled us to avoid falling into the trap of reporting on gender representativity and to inquire into the structural status quo that makes it difficult for women to have influence. Our findings revealed structural deficits that limit women's empowerment even in a country where matrilineal family structures are the norm. This is an example of the metaphor of Global North–Global South power relations at a community and district level. We met and spoke to powerful women at the community level who are continually speaking to power and engaging authorities but still struggling to find the agency needed to shift the system.

Contradictions we encountered in this work led us to ask what was missing from our analysis even though we expanded our system of analysis. We found that it was an analysis of how patriarchal systems of governance and traditional cultural systems overpower community systems of matrilineal leadership. The process tracing revealed that the way the project approached the empowerment of women was simplistic and examined empowering the individual with training but not the gender contradictions in the system in which they need to fight and have influence. One project alone cannot handle such systemic problems, and this left us with the question: What role do Global South and Global North organizations need to play to enable this level of coordination?

In this story, there was *complementarity* within the team when it came to working with these structural inequalities, with John and Mutizwa bringing their experience of working with feminist concepts of power into the evaluation from ACCRA, Jane her background in gender work and Chimwemwe his embedded understanding of gender in cultural and political systems in Malawi. We were a split between Northern partners, or the international consultants in the evaluation, and here we include Mutizwa, because he and John were the team players that normally would be seen to be holding the power of approach, and Jane and Chimwemwe as the regional players gathering evidence. The team did not accept this power dynamic that we encounter so often when Global South and Global North evaluators collaborate. The norm is that those who hold the approach or the theoretical position of an evaluation have more control or power over

analysis and results, but rather than certain knowledge and roles being given more power than others, we were able to appreciate our different knowledge as equally valuable and necessary for untangling this complex space.

The Malawi case study works with the concept of the Global North and Global South divide as a metaphor for power relations (both historical and current power relationship) and how these play out in the practice of evaluation rather than a more simplistic view of the Global North–Global South dynamic as a geographic description of unequal power relations. It also suggests that systemic change evaluation designs should consider multilayered complementarities and contradictions.

Joint Reflections from Our Community of Praxis

With its focus on addressing the manifestation of metaphorical Global North–Global South power relations at the local and national levels and tracking the potential for transformational shifts in governance practices, the team was able to deepen the design and evaluation approaches that we had explored in our previous assignments in Pakistan and in the ACCRA evaluation. This involved expanding the process tracing approach that Oxfam GB had developed, enabling us to investigate three interlinked levels of governance but centred on the experience of beneficiaries, particularly women.

Centring power in this evaluation also involved *grounding the feminist concept of power* in the way we evaluated how women were benefiting, enabling us to move beyond Oxfam's simplistic concept of women's empowerment to consider some of the *underlying cultural and systemic contradictions*.

These approaches were also reflected in the *further development of our internal praxis*. We were able to build particularly on *new-found complementarities* within the team, enabling us to appreciate different knowledge as equally valuable and necessary for untangling the complexities of power and governance dynamics within an expanded system of analysis. Mutizwa, as team lead, played a vital role in encouraging this equitable co-ownership of process.

EBA (Sweden): Working with a Global North Country Seeking to Work in Solidarity with the Global South

Our final story takes us to Sweden, where we recently completed an evaluation of long-term climate change investments on the global stage involving four of the five of us – Jane, John, Mehjabeen and Mutizwa (Colvin et al. 2020). Our evaluation of interactions between different layers of governance embraces the contribution of multilateral investments alongside bilateral investments in the context of complex multilevel landscapes, including North–South relationships.

Evaluators' Story

In early 2020, we completed an ex post evaluation of the Swedish Climate Change Initiative (CCI). CCI, executed from 2009 to 2012, was a demonstration of Sweden's commitment to fast-track climate financing for climate change adaptation and mitigation.

Sweden invested 4 billion Swedish krona through CCI, of which 72 per cent was distributed across 17 multilateral funds, 15 per cent to five countries with low adaptive capacity and high vulnerability to climate change and 13 per cent to two regional investments in Africa and Asia. As evaluators, our two main questions were: Has the CCI contributed to sustainable climate change adaptation and mitigation in poor countries and, if so, why, in what ways, and to what extent? What lessons from the CCI can inform Sweden's climate aid today?

This brief challenged us to develop an evaluation design that addressed the complexities inherent in a portfolio of this size. It involved diverse histories, contexts, investments, governance systems and programmes. We also had the challenge of conducting a contribution analysis over a decade. We drew on our accumulated evaluation experience and deepened our evaluation design praxis.

Three features of this evaluation stood out for us and frame this story:

- Our client – the Swedish Expert Group for Aid Studies (EBA) – accurately perceived us as a niche community of practice that negotiates effectively between the Global South and Global North. We were chosen based on the quality of our proposal, which for us, knitted everything we had learned together. This opened the door for us to a discerning client, one that is conscious of its place as a member of the Global North in the development context.

- The client did something unusual: asked for a 10-year sustainability and contribution analysis. Beyond this request, the terms of reference were open and unprescriptive.
- Based on research within the Swedish system, our client understood the tensions between accountability and learning. One of their own research papers asked why few intended users read and learn from evaluation reports, noting that the greatest learning happened for consultants and not for the intended audience (Reinertsen, Bjørkdahl and McNeill 2017). Also, consultants often developed recommendations for, rather than co-developing recommendations with, the intended users. We responded by offering our learning about this tension, centring a principles-based, utilization-focused approach with learning at the core.

Our evaluation challenges included the following:

- We had to bridge the relationship between Sweden's investment in its multilateral portfolio, involving board-level, global influencing and ground realities in different countries, referred to as 'the nuts and bolts of adaptation'. There was also a regional component to this multilevel relationship.
- Mirroring this, we had to bridge the relationship between the MFA, which is responsible for managing the multilateral portfolio, and Sida, which manages the bilateral portfolio.
- We had to research and construct 10-year contribution stories across this multilevel landscape, spanning the Global North and Global South. This included how to surface the significance and influence of CCI's principles-based approach, carefully constructed around the work of an international commission – the Commission on Climate Change and Development (CCCCD) – focusing on climate change adaptation. Within this, we would need to determine Sweden's leadership role, style and particular qualities.

Another evaluation realm required understanding the values and risks of this large surge in funding, making explicit the choices and understanding the nature of negotiations among its own key stakeholders that Sweden encountered in assembling a coherent investment portfolio for this funding surge.

Each of the challenges above entailed contradictions. We were able to draw on our joint experience as a community of praxis, including our learning-based design experience, and a proposed set of principles for

co-designing the process. With a utilization-focused learning approach, this would involve emergence and the need for adaptive management. We proposed a co-design approach with the Evaluation Reference Group (ERG), and our client agreed. The ERG then became our holding framework, offering tensile strength and complementarities to match the difficulty of the task. It allowed us to draw from a diverse evaluation toolkit compiled from our varied, complementary and sometimes contradictory skills.

Immediately, we identified a place of entry by grounding our evaluative research in bilateral country case studies. We visited Mali and Cambodia for deep research and developed a bilateral portfolio analysis. We found, for example, that national climate funds were particularly important for change at the national level; in Mali, these have strengthened national ownership of climate practices to include subnational and local systems through operationalization of the Mali decentralization policy.

We also moved up a level to evaluate a challenging portfolio of CCI investments for regional Africa with investment at a whole-continent level, including transboundary challenges of shared resources. In this process, we developed tools to grasp and understand such a large portfolio of scale, sectors and interlinking pathways.

Simultaneously, we undertook two initial multilateral case studies – of the Global Fund for Disaster Reduction and Recovery and the Forest Investment Program – to begin to analyse the 70 per cent of the CCI portfolio with a multilateral focus, worth 2.9 billion krona. This allowed us to test a methodology for developing decadal process tracing and contribution analysis stories across global funds.

Finally, in a further learning loop with the ERG, we undertook a full portfolio analysis of 17 multilateral funds and programmes, drawing on four of these as in-depth case studies⁹ and the remainder as shallower rapid reviews.

In summary, a critical design insight of this evaluation, with its breadth and complexity and global-to-local scale, was that, in a reversal of investment size, we studied the smaller, national case studies first and only subsequently considered the larger, multilateral case studies. This reversal was aligned with a core theme of the CCCD, which highlighted the importance of adaptations at the local, contextual level, and it allowed us to begin

⁹ These included case studies of the Adaptation Fund and the Clean Technology Fund in addition to the Global Fund for Disaster Reduction and Recovery and the Forest Investment Program case studies.

by studying local context and history as framing conditions for successful adaptation.

After observing maladaptation and successful adaptation at this smaller scale, our ability to enquire critically as to what the multilateral funds were doing and the possible repercussions of their decisions was enhanced.

Having analysed Sweden's quality, style and culture of leadership at the national and regional levels, we were better able to investigate Sweden's unique value addition at the global multilateral board level, adopting an approach that also stemmed from our commitment to understanding context and history.

Our major insight from this for transformational work is the value of centring on the local and working up and out from there. Impacts from climate change are felt at the local level, and that is where change must happen. All the other levels must revolve around this. To be equitable in good adaptation, we surfaced the imperative to centre the local first.

Joint Reflections from Our Community of Praxis

In this story, we can see an interweaving of all five of the themes explored in the previous stories and an emergence of all of our learning for transformation.

One strong theme is design for transformational learning. Here, a principled approach that included co-design with the ERG created adequate space over an 18-month period for ERG members – in particular those from the Ministry of Foreign Affairs (MFA) and the Swedish International Development Cooperation Agency (Sida) – to engage effectively with a complex set of questions and findings related to transformational design and practice for climate-resilient responses. An important feature of this learning process was that it was cross-organizational as well as individual (with a series of seminars that the MFA, Sida and EBA convened at the end of the assignment seeking to widen and extend this cross-organizational learning process).

The value of centring on the local and working up and out from there was also a significant aspect of our emerging methodology, which was enabled through the principles-based co-design approach and its affordances for emergence and adaptation in design. In line with the emphasis of the findings of the CCCD that local and contextual responsiveness is critical for climate adaptation, we were able to centre subaltern and Global South experiences through case studies in Cambodia, Ethiopia and Mali before attempting an analysis of CCI's multilateral portfolio – an important

reversal that enabled us to illuminate contradictions between global, multilateral approaches and different investment strategies at the national and local levels. Once again, we drew on our feminist analysis of power to inform these early design decisions, which also resonated with the client.

Centring the local also required drawing on our previous experiences of local engagement in multiple countries and how this starting point can lead to multiscaled analysis of complementarities and contradictions. Having Global South and Global North polarities in our team for this assignment, as well as our experience in bridging these, fostered the requisite internal praxis and holding framework for this demanding, global-to-local evaluation assignment.

Concluding Insights

This chapter has drawn on a conceptualization of transformation that foregrounds systemic change in what we originally framed as the Anthropocene context. Looking back, we are convinced that, in renaming the context, which shows meta-level external reflexivity and shapes our work, we embrace the concept of Capitalocene in place of Anthropocene. Capitalocene clarifies that it is not the whole of humanity that is responsible for the current crises; rather, capitalism is (Moore 2015). Fortunately, this realization does not change our story, because capitalism largely defines relationships between the Global South and Global North. Through a series of stories from our community of praxis, we have illuminated several themes that we see as critical to transformational design, with a primary focus on transformational design for evaluation of and for systemic transformation.

The first two of these themes, which are closely interwoven, are concerned with the *centrality of navigating power* in South–North *complementarities and contradictions*. In Pakistan, we worked with contradictions between expert and indigenous knowledge and the power dynamics that shaped these. In each of our evaluations, we applied feminist concepts of power in several ways – in analysing power relations and their transformation, in giving space to multiple voices and perspectives and in informing the complementarities and contradictions in our internal praxis as a team. In the Malawi and EBA evaluations, we centred local and women’s voices and experiences within much larger systems of analysis to reveal complementarities and contradictions within multilevel governance systems and their design.

Analysing and navigating power in South–North complementarities and contradictions requires a good understanding of *history and context*.

In the ACCRA and EBA evaluations, we used learning history as a praxis for building and narrating historical context, and through the EBA evaluation, we gained critical insight into the importance of centring the local for transformation praxis, in retrospect drawing on this methodologically in the emergent evaluation design as we sought to navigate a complex portfolio of multilevel governance shaped by multilateral and bilateral investments.

We also experienced the value of *good internal praxis* within the team, enabling us to act as skilled facilitators of transformational (evaluation) processes. Underpinned by North–South complementarities and contradictions, internal praxis in some stories called on us to face our fears and vulnerabilities and to navigate these together, guided by feminist concepts of *power with* and *power to*. Internal praxis also invited us to work with mirroring as a way to understand and engage with (hidden) dynamics in the external system of interest and helped us rebalance the significance of multiple voices in the evaluation process, centring around the local and subaltern while also recognizing the value of bridging multiple levels and between hegemonic and subaltern knowledges and ways of knowing.

A core aspect of praxis that comes to the fore in these stories, with particular relevance for transformational evaluation processes, is the design and facilitation of *adaptive and potentially transformational learning processes*. Although contradictions of ownership in the Malawi story prevented us from effectively positioning learning, adaptive learning was a core feature of the other three stories and was perhaps most effective in the EBA evaluation. Here, the introduction and expansion of effective cross-institutional learning depended on three factors – the openness and receptivity of the client, our own individual and collaborative skills as facilitation practitioners and our joint ability with the client to create an effective institutional holding framework.

In a COVID world, where the pandemic is already exacerbating poverty and inequity, the five themes presented here as underpinning design for transformation as systemic change become even more central to global-to-local responses. The shocks and uncertainties of the pandemic and the need for an accelerated global response to climate change require engagement with the types of complementarities and contradictions between the Global North and the Global South that we have been recounting in this chapter.

Now, more than ever, our story reminds us of the crucial role of good internal praxis as we are called upon to face our fears and vulnerabilities and to navigate these together, embracing diversity and power differentials and guided by feminist concepts of *power with* and *power to*. By honing relevant

skills, assumptions and framings, leadership teams, collaboratives and social movements seeking to address global-to-local problems such as the pandemic and climate change will be better equipped to navigate power in South–North complementarities and contradictions.

When the evaluation community is called upon to look back on these times and reflect on how we collectively addressed these global problems, our learning also shows us that an understanding of history and context must play a central role in making sense of complexity. Furthermore, this calls upon evaluators to become researchers and facilitators, not only to delve deeply into knowledge systems, but also to centre the design and facilitation of adaptive and potentially transformational learning processes at the heart of evaluation praxis in the service of transformation.

A final conclusion is that, in all of the above, it remains important to continually weave theory and practice together. In this chapter, we have sought to demonstrate how we practice this as a community of networked professionals. We have explored transformation within our own work together as a niche community of praxis, as well as in our evaluation and design work with stakeholders and partners. We have shared our story of growth, failure and maturing through our exploratory work and trust that this will resonate for others. In this spirit, we offer it to anyone who identifies as being on a similar journey or would like to start one. In the new world we now inhabit, our stories of growth, failure and maturing become valuable tools and perspectives for the journey ahead, where our collective abilities to reflect on our own humanity in the context of complex global situations must take centre stage.

Acknowledgements

This chapter is the result of a collaboration in which leadership continually shifted between different members of our community of praxis. We have therefore listed authors in alphabetical order, intended to denote a non-hierarchical author contribution.

We would like to thank the British government and the Swiss Agency for Development Cooperation for funding our participation in the IDEAS 2019 conference. Without this assistance, we would not have been able to participate.

References

- Arendt, Hannah. 1969. *On violence*. San Diego: A Harvest Book.
- Boulding, Kenneth E. 1990. *Three faces of power*. Newbury Park, CA: Sage Publications.
- Bradbury, Hilary, George Roth and Margaret Gearty. 2015. The practice of learning history: Local and open system approaches. In: *The SAGE handbook of action research*, 3rd ed., (ed.) H. Bradbury, 17–30. London: Sage Publications. <https://dx.doi.org/10.4135/9781473921290>
- Climate Investment Funds. 2019. *Transformational change in the CIF: Summary of findings from an independent evaluation and evidence synthesis*. Washington, DC: Climate Investment Funds.
- Colvin, John, and Mehjabeen Abidi-Habib. 2013. Learning about transformational design in Pakistan's climate compatible development space. Paper presented at Proceedings of Transformation in a Changing Climate, 19–21 June, Oslo, Norway. Oslo: University of Oslo.
- Colvin, John, Mutizwa Mukute, Mehjabeen Abidi Habib et al. 2020. Evaluation of the Swedish Climate Change Initiative, 2009–2012, EBA Rapport 2020:02. Stockholm: Expertgruppen för biståndsanalys.
- Engeström, Yrjo. 2001. Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education Work* 14 (1): 133–156. <https://doi.org/10.1080/13639080020028747>
- Fazey, Ioan, Niko Schöpke, Guido Caniglia et al. 2018. Ten essentials for action-oriented and second order energy transitions, transformations and climate change research. *Energy Research and Social Science* 40: 54–70.
- Gaztambide-Fernandez, Ruben A. 2012. Decolonization and the pedagogy of solidarity. *Decolonization: Indigeneity, Education and Society* 1 (1): 41–67.
- Hawkins, Peter. 1993. *Shadow consultancy*. Bath: Bath Consulting Group.
- Heifetz, Ronald, and Donald L. Laurie. 1997. The work of leadership. *Harvard Business Review* 75 (1): 124–134.
- Karlberg, Michael. 2005. The power of discourse and the discourse of power: Pursuing peace through discourse intervention. *International Journal of Peace Studies* 10 (1): 1–23.
- Kubera, Paulina. 2019. Evaluation as reflective practice. *Problemy Zarządzania – Management Issues* 16 (6): 23–38. <https://tinyurl.com/ykvymh44>
- Levine, Simon, Eva Ludi and Lindsey Jones. 2011. *Rethinking support for adaptive capacity to climate change: The role of development interventions*. London, ODI.
- Moore, Jason. 2015. *Capitalism in the web of life: Ecology and the accumulation of capital*. London: Verso.
- Mukute, Mutizwa, Jane Burt, John Colvin et al. 2021. Impact evaluation of the 'Kutukula Umwini pa Chitukuko Cha m'maboma Ang'ono project: Enhancing communities' capacity in tracking development resources at local council level'. Oxford: Oxfam GB.

- Mukute, Mutizwa, John Colvin and Aristides Baloi. 2017. *Africa Climate Change Resilience Alliance (ACCRA) programme phase 2: Synthesis evaluation*. Oxford: Oxfam GB.
- Murove, Munyaradzi F. 2009. An African environmental ethic based on the concepts of ukama and ubuntu. In: *African ethics: An anthology of comparative and applied ethics*, (ed.) M. F. Murove, 63–84. Pietermaritzburg, South Africa: University of Kwazulu-Natal Press.
- Oxfam GB. 2016. *The future is a choice: The Oxfam framework and guidance for resilient development*. Oxford: Oxfam GB.
- Reinertsen, Hilde, Kristian Bjørkdahl and Desmond McNeill. 2017. Confronting the contradiction – an exploration into the dual purpose of accountability and learning in aid evaluation, EBA Rapport 2017:06. Stockholm: Expertgruppen för biståndsanalys.
- Roy, Joyashree, Petra Tschakert, Henri Waisman et al. 2018. Sustainable development, poverty eradication and reducing inequalities. In: *Global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*, (eds.) V. Masson-Delmotte, P. Zhai, H.-O. Pörtner et al. IPCC. <https://www.ipcc.ch/sr15/chapter/chapter-5/>
- STAP (Scientific and Technical Advisory Panel). 2017. Strengthening monitoring and evaluation of climate change adaptation: A STAP advisory document. Washington, DC: Global Environment Facility.



CHAPTER 7

Governance Pathways for the Greater Caribbean: Transformative Evaluation Principles

LENNISE J. C. BAPTISTE

Abstract. This chapter was developed from a thematic review of publicized statements of stakeholders of the Association of Caribbean States (2017) to understand their priorities for post-COVID-19 recovery in the region. The analysis showed that regional priorities were to improve health systems, ensure food security, improve transportation channels to access supply chains for medicine and food, develop partnerships to leverage economies of scale, preserve the environment and develop the economy. The COVID-19 pandemic has illuminated the gaps in governance systems that were designed to keep citizens safe and provide relief in times of crisis. In this chapter, the governance systems of countries in the Greater Caribbean were examined to identify how transformation change practices could help in the crisis management and recovery phases. Transformative evaluation practices and establishment of internal monitoring and evaluation systems were proposed to increase demand for evaluation to support decision-making and build an evaluation culture. Capacity building, strategic planning, policy development and use of information and communications technology were identified as transformation pathways for the region.

Introduction

Every country in the world is working to respond to a health crisis that has severely limited economic activity, food security, formal education processes, migration, citizen security, transportation within and between countries and, in some countries, confidence in government leadership. Worldwide, the COVID-19 pandemic has illuminated the gaps in governance systems that were designed to keep citizens safe and provide relief as governments have grappled with unbudgeted but necessary emergency expenditures in the health care sector. Established international systems for finance, trade, travel and communication were also disrupted, exposing the differences between countries and regions in the financial power needed to negotiate successfully for needed resources.

This chapter focuses on the Caribbean and describes the challenges that countries in the region face. Sustainability, development and systems thinking in national and regional governance systems are discussed in an examination of the Caribbean context. The Prague Declaration on Evaluation for Transformational Change provided guidance on evaluation for transformation practices¹.

Effect of COVID-19 in the Greater Caribbean

Many governments in the Greater Caribbean² are facing shrinking revenues within their countries and the dilemma of paying their debts to international financial agencies while also facing the additional costs of emergency funding from the same pool of agencies. Regional economies are 'already besieged by both climatic and economic shocks, including heavy indebtedness and high exposure to natural disasters'³, and many regional leaders have asked the United Nations Economic Commission for

¹ Adopted 4 October 2019 by the International Development Evaluation Association Global Assembly and the Third International Conference on Evaluating Environment and Development.

² The Greater Caribbean Zone of Co-operation was established in 1994 and consists of joint actions in the priority areas of the Association of Caribbean States: trade, sustainable tourism, transport and disaster risk reduction in recognition of the common geographic space that the states, countries and territories in the Caribbean Sea share.

³ Remarks by Alicia Bárcena, Executive Secretary ECLAC, at virtual meeting hosted for Caribbean heads of state and finance ministers and United Nations resident coordinators and agency representatives in the Caribbean, 29 April 2020.

Latin America and the Caribbean (ECLAC) to appeal to the international financial community on their behalf for better access to grants and concessional financing.

Described as the 'world's most trade and travel dependent region' (CARICOM 2020) and facing 'trebling unemployment, and halved government revenues due to COVID-19' in the Caribbean before the pandemic, a 'high degree of inequality, combined with the high levels of poverty, informality, lack of social protection and limited access to quality timely health care, explain the high social costs that the pandemic is having in the Region' (ECLAC and PAHO 2020). Caribbean countries had to respond to domestic challenges such as 'revenue and income losses, a drop in investment, rising unemployment, increased indigence and poverty, the failure of small and medium sized businesses, and challenges to the financial system'⁴. The region also faced the external challenges of 'near total shutdown of air and cruise travel...stress in related supply chains (agriculture, construction, hotels, restaurants)...contraction in larger economies...downturn in commodities prices...contraction of foreign direct investment (FDI) flows and remittances...disruption in transportation and global supply chains; risk aversion for external investors and financial turbulence, and restrictions on foreign exchange availability' (ECLAC 2020a).

Levaggi (2020) suggested that the COVID-19 pandemic presented an opportune moment to address 'the crisis of regionalism manifested in the limited regional responses to technical issues to the collective challenges faced, and the limitation of state resources'. He recommended the promotion of 'good practices in bilateral and multilateral co-operation in the region' by examining successful cases of cooperation during this pandemic. He also recommended regional monitoring and evaluation (M&E) of the consequences of COVID-19 and focusing on the 'deepening humanitarian crises...the destabilization of democratic institutions and the impact on human rights'. These recommendations highlight the need for linked regional and national M&E systems to provide valid, credible information about the common challenges that countries face. Among these challenges are governing with limited national resources; negotiating regional, bilateral and multilateral collaboration; addressing humanitarian crises and navigating operations within regional democratic institutions. Underpinning these recommendations is advocacy for changes in national governance systems, changes in regional and international systems for trade and cooperation

⁴ Remarks by Alicia Bárcena, 29 April 2020.

and changes in how regionalism is valued and leveraged to derive the greatest benefits for regional collective development by participating in international systems as a bloc.

Spanish, French, Dutch and English are spoken in the Greater Caribbean, which consists of 37 countries touched by the Caribbean Sea with diversity in topography, culture, governance structure and development status. The Association of Caribbean States (ACS) facilitates 'consultation, cooperation and concerted action' among countries in its membership. The principal organ is the Ministerial Council, comprising representatives from the member states working with four special committees to develop joint actions for cooperation in four priority areas: *trade development and external economic relations*, focusing on shared economic space, regional trade statistics, studies and training programmes in trade negotiations; *sustainable tourism*, establishing a sustainable tourism zone of the Caribbean; *transportation*, regional cooperation for transport and connectivity; and *disaster risk reduction*, regional disaster planning, relief, prevention and risk mitigation.

The ACS meetings include members, associate members, observer countries, observer organizations and social partners⁵ that provide guidance for the decision-making process. The critical role of the ACS was

⁵ *Members:* Antigua and Barbuda, Bahamas, Barbados, Belize, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, St Kitts and Nevis, St Lucia, St Vincent and the Grenadines, Suriname, Trinidad and Tobago, Venezuela.

Associate members: Aruba, British Virgin Islands, Curacao, France on behalf of French Guiana and Saint Barthelemy, Guadeloupe, Martinique, Saint Martin, Sint Maarten and The Netherlands Antilles on behalf of Saba and Sint Eustatius.

Observer countries: Argentina, Belarus, Bolivia, Brazil, Canada, Chile, Ecuador, Egypt, Finland, India, Italy, Japan, Kazakhstan, Kingdom of Netherlands, Korea, Morocco, Peru, Palestine, Russian Federation, Serbia, Slovenia, Spain, Saudi Arabia, Turkey, Ukraine, United Kingdom, Uruguay, the United Arab Emirates.

Observer organizations: The Caribbean Community Secretariat, the Latin American Economic System, the Central American Integration System and the Permanent Secretariat of the General Agreement on Central American Economic Integration were declared founding observers of the ACS in 1994. ECLAC, Caribbean Tourism Organisation, Bolivarian Alliance for the Peoples of Our America – People's Trade Treaty, Central American Economic Integration Bank, European Union, International Organization for Migration.

Social partners: The Antilles-French Guiana Regional Centre of the National Institute of Agronomical Research, Association of Caribbean Universities and Research Institutes, Association of Caribbean University, Research and Institutional Libraries, Caribbean Association of Industry and Commerce, Caribbean Conservation Association, Caribbean Medical Association, Caribbean Shipping Association,

heightened in the regional response to the COVID 19 pandemic, when supply chains for food and essential health supplies were interrupted, and transportation (shipping especially) routes had to be reorganized because of closed borders within and outside the region. The role was described as follows:

The ACS can complement national efforts through promoting the sharing of best practices in a way that is targeted and meaningful to all sectors under its purview, thereby arming Member States with specific actions to deal with the negative consequences of the novel coronavirus in the short, medium, and long-term... The ACS has leveraged its Membership and engaged partners at all levels – the national, regional, hemispheric, and international – to facilitate information sharing and gathering. The compilation and analysis of information can prove a good resource to Member States, to gain a better appreciation of the existing synergies among different focal areas and sectors (Persad 2020).

Methodology

This chapter seeks to highlight how transformative change practices could be implemented to improve the performance of national and regional governance systems in the Greater Caribbean for the management and recovery phases of the COVID-19 crisis by increasing the demand for evaluation. The methodology comprised a thematic review of statements that regional ACS stakeholders published to identify how evaluation was integrated into their recovery and explain the complexity of national and regional governance systems regarding the shared economic space, preserving the environmental integrity of the Caribbean Sea, the promotion of sustainable development and the embrace of regionalism.

Findings

Role for Evaluation

There is increased understanding about the utility of data and evaluation findings at the regional and national levels because of the unanticipated

immediate need for a volume of information for decision-making across sectors due to the COVID-19 crisis.

Regional Thinking

The ACS is trusted to provide guidance on how to access resources (medical supplies and equipment being the immediate need) and facilitate cooperative relationships within and outside the region. Members were committed to leaving no citizen, city or country in the region behind. The interdependence of member states for food security and protection of the vulnerable and the importance of having an inventory of regional resources and a system (a humanitarian corridor⁶) to facilitate movement of critical resources across borders were acknowledged. There is need for more communication and cooperation between the regional organizations and regional-level monitoring of public policies to identify best practices to address the current threat. The transportation sector was described as fractured and needing collaboration and cooperation to bolster regional transportation.

Sustainable Development

Stakeholders identified regional priorities as improving health care systems and ensuring equitable access to services; ensuring food security; having multiple open channels of transportation to access supply chains for medicine, equipment, food and agriculture, and infrastructure; developing partnerships to leverage economies of scale and learn from best practices from implementation of development initiatives; preserving the environment and economic development, with opportunities to reduce debt to international financial institutions.

Complexity in Governance Affecting Management and Recovery

Limited use of technology by governments has reduced the efficiency and effectiveness of service delivery to citizens. Factors such as underdeveloped communication channels between governments and citizens, lack of technical capacity of citizens in some sectors and inequitable access to health care and education are impeding recovery efforts. Nationally, the

⁶ ACS Secretary General, Dr. June Soomer, 7 April 2020, at the ACS Founding Observer Organisations Exchange Initiatives to Coordinate COVID-19 Response.

siloed approach taken to governance, lack of policy harmonization between sectors, lack of policy coordination within sectors and scarce financial resources contribute to governance challenges. Harmonization and coordination between regional and national policies is lacking.

National and Regional Governance Systems in the Caribbean

In 2018, three actions were identified that could improve the delivery of services by state institutions in Latin America and the Caribbean (ECLAC 2018, 141). First, state institutions must 'be oriented towards building states that are more trustworthy and based on stronger guarantees of the rule of law and that promote fair competition'. Second, states were advised to strengthen administrative capacities by adopting easily adaptable, effective, efficient bureaucratic procedures; recruiting and attracting competent, qualified civil servants; improving the coordination and upgrading of management policies and long-term strategic plans; improving co-ordination between levels of government and across sectors and improving the response to national, regional and international factors that affect governance. Third, states were pressed to develop capacities for the effective use of information and communication technology (ICT) as a priority, to strengthen regional democratic processes and open government channels of participation.

Managing for development results was proffered as suitable approach for Latin American and Caribbean countries because it prioritizes strategic foresight, facilitates improvement of public management by focusing on achievement of measurable results and strengthens M&E processes. The five components of the approach are results-oriented planning, results-based budgeting, public financial management, project and programme management and M&E systems. 'LAC [Latin American and Caribbean] governments have made less progress in the areas of evaluation of spending effectiveness, aligning incentives to achieve institutional objectives and implementing evaluation systems' (ECLAC 2018, 162).

Systems Thinking and Governance in the Caribbean

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' (Magro and Van den Berg 2019, 144). The governance systems in Latin American and Caribbean countries comprise units, departments and ministries that undertake one or more of the processes in the five components of the managing for development results approach. Each of those sections of governments acts as a mini system within the governance system, and those mini systems are a mix of hard and soft components⁷. Systems are defined according to identifiable, agreed-upon boundaries; identified roles, responsibilities and unit mandates; relationships between units with established protocols and identified positive and negative feedback loops. Without the required system definition, the overall governance system would be mired in confusion, with competition for power and authority, resources and stakeholder alliances as the people in those systems (units) work continuously to maintain their relevance and essentiality.

For a transformational recovery in the Greater Caribbean, the complexity within individual and interrelating systems in national and regional governance systems must be identified so that the issues can be addressed and not prevent countries from making progress⁸. Feinstein (2019, 20) suggested that governments should change their evaluation focus from 'projects and programmes to strategies and policies' to obtain an accurate answer to the questions: Are we achieving our strategic intent? Are our systems effectively meeting the needs of our people? Is there coherence in the services offered to citizens?

The resilience of countries in the Greater Caribbean is being tested with the management of and recovery from the COVID-19 pandemic. Regional leaders and citizens have acknowledged the need for national

⁷ 'Hard systems – concrete components; soft systems – as legal, institutional, religious, cultural and art systems that may involve concrete components but overall sets of values, beliefs, principles, rules etc.; mixed systems – composed of soft and hard units/systems' (Magro and Van den Berg 2019, 145).

⁸ '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' (Magro and Van den Berg 2019, 146).

and regional governance systems to change, but the scope and depth of change has not been ascertained. What to do? Where will the funding come from? How much additional debt can be taken on without the total collapse of regional governments? Should the 2030 Agenda for Sustainable Development be abandoned? Regional states are struggling to answer these questions and reset their development priorities in the context of dwindling financial resources. Systems thinking can illuminate the components that influence change and the synergies, redundancies, strengths and weaknesses that can inform decision-making (Hargreaves 2010).

United Nations Secretary-General António Guterres recommended that states turn their recovery into opportunities to 'do things right for the future...and steer our world on a more sustainable path'⁹. Any change process for governance systems in the Greater Caribbean must embrace the complexity of interacting systems of new digital technologies, global knowledge networks, environmentally aware consumers, new technological options, scientific progress and sustainable development strategies. This complexity will also be present in the systems of regional institutions that support engagement with individual countries and with external agencies on behalf of countries. 'Understanding this co-evolution is the basis for expediting the change towards sustainability... Equality and sustainability can only be placed at the centre of the development pattern if social compacts are constructed to make this possible, because development is ultimately a political issue' (ECLAC 2020b).

Recognizing the Value of Evaluation

In attempting to envision how evaluations can be reframed and conducted in this time of crisis, consideration must be given to the various perspectives on what evaluators do and what is valuable about monitoring, evaluation and learning (MEL) processes. There are some dominant perceptions in the region among persons at different levels of government such as evaluation reports are for the funders, they highlight successes and failures and inform decisions about whether to continue funding, and governments cannot use MEL processes without more funding to establish M&E systems (equipment, software, protocols) and provide the required retraining of personnel.

⁹ Use COVID lessons to 'do things right' for the future, urges UN chief – from online remarks to the World Conference of Speakers of Parliament (see <https://news.un.org/en/story/2020/08/1070652>).

For government leaders and senior public servants, buy-in regarding the value of evaluation is still not enough to realize transformational changes needed for managed and delivered citizen-centric services. Governments commission evaluations of donor-funded government initiatives to satisfy donor requirements, and recommendations are often not applied to the national development process nor are findings used to revise established government practices. By maintaining this perspective, country leaders are redirecting responsibility for sustainable development and improvement of national processes to actors outside of their countries.

Over the last two decades, many persons in government and civil society have been exposed to project-related M&E training for progress reports and final reports. There has been an increase in the number of trained regional evaluation professionals who have progressed beyond that training and have studied evaluation approaches at accredited institutions outside the region, although the regional options have increased. Recruitment of regional professionals into evaluation teams for donor-funded regional projects has increased. Challenges such as access to stakeholders, conflation of the process of project evaluation with those of impact and outcome evaluations, and lack of financing of M&E nationally and regionally continue to diminish the potential of regional evaluation activities to provide comprehensive MEL information and guide strategic planning. For MEL activities to influence the transformation of countries, leadership from the highest levels of government and regional organizations, must lead the charge to build an evaluation culture in which accountability is valued (Baptiste et al. 2019). Wiltshire (2015) advised that, to strengthen governance in the region, data collection and monitoring systems must be improved.

There are two major obstacles to the establishment and use of national evaluation systems within the Caribbean. First, although the value of M&E data is increasing, and there are data collection and reporting systems that can be built up and human capacity that can be repurposed, country leaders have not been enthusiastic about implementing internal M&E systems. Second, the failure to commit budgetary allocations to evaluation processes signals that MEL is not important or essential for good governance and decision-making. For many governments, the value of conducting evaluation activities can become lost among competing priorities and demands for financial resources, for what may be perceived as more important and urgent, with more tangible, visible results. Government staff and stakeholders may resist implementation of evaluations because of a lack of understanding and agreement about evaluation criteria, feelings of loss of power and control over the process and the use of outcomes (Taut

and Brauns 2003). Government stakeholders may also fear that evaluation reports will illuminate shortcomings of government processes, low technical competence levels of personnel and inability to conduct project activities. Stakeholders often fear that evaluation reports will be limited to what is measurable in the results frameworks and that critical contextual data about a programme's response to unintended beneficiaries, unintended outcomes and its non-quantifiable impact would be lost (Baptiste and Moss 2017).

There are also conditions (complexities) in national contexts that may not readily support accountability, transparency and use of MEL at the government level, including a lag in high-level decision-making due to partisanship from feuding political parties, skewed choices of development projects or project implementation due to the influence of financial contributors to political parties, non-citizen-centric policies, ambiguous procurement rules, weak law enforcement, short election cycles and discontinued development initiatives when governments and development priorities change. According to Matera and de Lourdes Despradel (2020), the ongoing challenges of corruption, public insecurity, organized crime and institutional weakness, as well as the climate change patterns of stronger and more frequent hurricanes and drought in Central America, continue to plague the Greater Caribbean. These authors also suggested that the competing ideologies of the United States, China and Russia have affected good governance, transparency and effective security in the region.

Many of the ACS member states have middle-income status, but high levels of debt owed to international financial institutions have hampered their development in health, education, infrastructure, administration of justice, social protection, food and nutrition security and other areas (Wiltshire 2015). Thus, in this time of crisis, regional governments must prioritize how they use their available resources and those that they are able to access. Before this crisis, Wiltshire (2015) suggested that regional priorities should be establishment of a framework for strengthening governance, improving data and monitoring systems for building effective partnerships and implementing and monitoring the Sustainable Development Goals (SDGs), which Wiltshire (2015) described as discrete goals with proposed indicators for monitoring. She advised that, if the SDGs were not 'approached as interconnected and interdependent', the region would not be able to slow or reverse the negative development trends. She recommended that regional leaders prioritize implementation of the SDGs that would 'strengthen economic performance, promote inclusive and transparent governance, support gender equality and sustainable development, and promote beneficial engagement with the global economy' (Wiltshire 2015, 9).

In 2020, responses to the COVID-19 pandemic has vividly illustrated the interconnectedness of governance problems, and comprehensive solutions are needed that encompass a variety of inputs from different sectors, meaning that the SDGs should be implemented using an integrated approach. 'Part of the reason that the world is not yet on track to realise the SDGs, is that policies and plans, as well as MEL efforts, have often failed to recognize the systemic nature of the SDGs' (Ofir et al. 2019).

Learning from evaluation can also contribute to the commitment of ACS member states and embrace of regionalism as a vision and a value, not only for this recovery process, but also for the long term. Evaluation findings from regional joint ventures can illustrate that the benefits of collaboration can benefit the resilience and development of the region as a whole, as well as for individual member states.

Transformational Change and Transformative Evaluation

The changes proposed to achieve transformation in national and regional governance identify actions that can be undertaken to 'do things differently' for recovery and continued development after the COVID-19 crisis. The following quotation provided further clarity about transformation and transformational change and the essential element of sustainability over time.

Transformation refers to change that is radical, revolutionary – whether in individuals, institutions, societies, countries, (eco)systems, or the planet as a whole... In the change spectrum it is at the other end opposite incremental change, although many incremental changes can – and often do – lead to transformation... Transformational change is the process whereby positive development results are achieved and sustained over time by institutionalizing policies, programmes and projects within national strategies (Ofir 2018).

Transformative evaluation is described as 'a branch of programme evaluation where social justice is the primary principle guiding an evaluator's work' (Bolinson, Mertens and Engineers Without Borders Canada, 6). This approach seeks to unearth the varied perspectives held by stakeholders, which emanate from their experiences, and it is ideally suited when highlighting marginalization, privilege, oppression, discrimination, inequality, power differences and inequity in the evaluation context using quantitative and qualitative methods. The evaluator must seek to build trust with

stakeholders so that they are confident in using the findings to address human rights and social, economic and environmental justice. The evaluator must also address the intersectionality of the culturally responsive, feminist, equity-focused and indigenous theories if they are relevant to understanding and interpreting the dynamics in the evaluation context.

Pathways to Better Governance Systems

In this section, four pathways are proposed to transform national and regional governance systems in the Greater Caribbean. These are not new pathways, but activities are proposed that, if implemented, have the potential to improve results for national and regional governance.

A Three-Tier Education Drive for National Governments and Regional Agencies

The purpose of this drive is to begin building an evaluation culture by increasing awareness of the value and utility of evaluation practices and findings, using relevant examples from the governance system. The engagement will be designed to reduce fear of participating in evaluation by illustrating that its value and utility is not just for reporting to donor agencies, but that it can also yield critical information about the progress of national and regional development.

Initial transformation is expected at the individual level as people begin to think evaluatively. Then, as they work collectively in their communities (units, departments, ministries, agencies) to integrate evaluation activities into everyday practice, those synergies will transform their communities.

The content will be organized to meet the needs of each of the three tiers. Tier 1 will comprise government leaders, national and regional technical advisory teams, agency heads and senior public servants in the highest grades with significant responsibility for policy (development, implementation, adherence, revision). The content will focus on evaluation for decision-making such as policy and strategy evaluation. Participation and buy-in of leadership will be necessary to emphasize that evaluation is valued within the governance systems. An important exercise will be a retrospective alignment of the SDGs with government projects over the past 15 years to identify the development focus nationally and across the region. A national follow-up could be harmonized collaboration of ministries and departments and of countries to procure and use resources more efficiently and effectively. A regional follow-up could be revision, development and

implementation of policies and strategies to facilitate achievement of the ACS objectives (enhance economic space, preserve the environmental integrity of the Caribbean Sea, promote sustainable development).

Tier 2 will comprise public servants who provide business and policy support, such as executive assistants, for specialist services in the areas of human resources, finance, information technology and communication. The content for Tier 2 will focus on the value of evaluation practice for doing things right and doing the right things, with the aim of providing feedback about operations at the project and programme levels. Tier 3 will comprise administrative staff, and the training content will focus on why data are needed, the importance of being accurate and how their roles contribute to realizing the vision of the unit, agency, ministry and government as a whole.

An important aspect of this initiative will be to dispel the myth that evaluation will focus on criticizing governments, so marketing evaluation as learning will be essential. Emphasis should also be placed on identifying and leveraging established data collection and management systems within the governance system that could complement an internal M&E system. The aim is to establish multidisciplinary working groups across different levels of government and across agencies who can act as coaches to support the work of an internal M&E system.

Transformative evaluation practices include working in partnership, exploring power relations, promoting inclusiveness and sharing responsibility for results¹⁰. As with any change process, resistance to change, defensiveness and assigning blame should be expected. The expected results, because of the increased information flow, are that roles will be established and defined to structure the internal M&E systems, and that the value of accountability and transparency will increase. The real-time evaluation approach¹¹ in training content could also be included to provide timely information for decision makers.

Strategic Planning

Citizens' experience with public services is a key determinant of satisfaction and trust levels in governments... Better understanding citizens' needs,

¹⁰ Prague Declaration on Evaluation for Transformational Change strategies 2, 3 and 9 (see <https://tinyurl.com/bcm294k7> and chapter 18 of this volume).

¹¹ 'Looks at the likely outcomes of current policies, not simply keeping track of whether targets are being met...all actors believe it can contribute to improving the ongoing response and unlock operational bottlenecks' (Polastro 2014).

experiences and preferences can result in better targeted services, including for underserved populations often at little extra cost (OECD 2017).

A non-negotiable stance against inequality and poverty is essential for making government measures more redistributive, particularly with regard to public finances, while supporting a shift towards a more inclusive structural change (ECLAC 2020b, 218).

The big push for sustainability provides an opportunity to build a new style of development based on a new equation between the State, market, society and the environment, which is, in essence, the key aspiration of the 2030 Agenda for Sustainable Development (ECLAC 2020b, 225).

These three statements established the focus for the kind of strategic planning aimed at transformation of the governance system. The statement from the Organisation for Economic Co-operation and Development (OECD) identified the first step: governments (leadership and all levels of public servants) must accept that their mission must be satisfying the citizenry with delivery of public services to increase citizen trust in public sector practices. The second step, a needs assessment, is essential to capture citizens' experiences and preferences so that services can become more citizen-centric, and a greater attempt must be made to understand the needs of the underserved portions of the population.

For the recovery period, regional governments will not have the time and financial resources to undertake new and comprehensive data collection and analysis across communities and government agencies. Thus, a meta-analysis and synthesis of findings and recommendations from regional studies on governance that organizations such as ECLAC, OECD and the Inter-American Development Bank have undertaken in the last 10 to 15 years could be conducted as a starting point. For systemic, transformational change, it is time to address the weaknesses that have been identified and repeated in several reports over the years, including slow rates of addressing institutional weaknesses such as underdeveloped administrative capacities, onerous bureaucracy and weak coordination and cooperation practices that undermine policy adherence and the rule of law, making public service operations more efficient and addressing corrupt practices.

Considering ECLAC's call to adopt a 'non-negotiable stance against inequality and poverty', the transformation challenge is for governments to regard their citizens as worthy of equitable access to economic and job growth, better quality of life, citizen security, gender equality, health services and education. Strategic planning must lead to results such as, but not limited to:

- Fewer interactions with different government personnel and visits to different government ministries to complete a transaction¹².
- Absence of long lines outside government buildings that begin hours before the scheduled opening – this usually happens because of a sudden change in policy or requirement for citizens doing business or because of reduced opening hours to address failing infrastructure.
- Shorter processing times for transactions vital to access services such as applications for national identification card, birth certificate, driver's license, passport.
- Publicized criteria and information for applications and completion of transactions (this could change based on information from security guards or frontline staff at government departments).
- General customer service training for all staff and as part of the orientation of new hires to learn how to satisfy internal and external customers.
- Accessible buildings and services for differently abled persons – trained personnel assigned to each ministry to work with frontline staff and other staff as needed.
- Decentralized public services and the use of community resources to increase economic activity within and across countries.
- Establishment of partnerships with tertiary educational institutions to address labour skills gaps in the population and increase the number of employable citizens or citizens who can generate their own income from self-employment.
- Alignment of governance plans with the SDGs and use of the United Nations 2030 framework to build a national results framework with a special focus on new areas of job creation needed to preserve the environment.
- Adoption of a whole-of-government approach emphasizing cooperation and coordination, from plan development to implementation and evaluation, to reduce duplication, inefficient use of resources and bureaucratic conflicts that hinder resolution (OECD, CAF and ECLAC 2018, 155).

¹² Too often, citizens must demonstrate that the previous 'steps' were taken, because information (mainly documents) was not sent from the last department visited, or information was lost in the transference from one department to the next.

Informed leadership, transparent governance, multidisciplinary expertise and funding are needed to engage in comprehensive strategic planning for the recovery of individual countries and the region.

Policy Development, Implementation, Coordination, Harmonization, Evaluation

'Weak co-ordination and co-operation practices at the political and administrative levels can greatly undermine efforts to achieve policy coherence' (OECD, CAF and ECLAC 2018, 165). To increase trust in public service, the desire to deliver citizen-centric services must be at the heart of policy choices and development to reduce inequality among citizens and increase citizen access. 'Inasmuch as governments use data as a strategic asset to boost public sector intelligence, they can improve services as well as their capability to develop sustainable and inclusive policies' (OECD, CAF and ECLAC 2018, 166). Inclusion is essential to eliminate siloed operations, increase coordination among government institutions and simplify service delivery and access. Effective coordination must comprise three components: focused communication and clarification about new processes and their specific objectives and expected results, final agreements from the redesign process informed by the opinions of relevant stakeholders and high-level political awareness and support for the new policies (OECD, CAF and ECLAC 2018, 169).

Use of Technology

The COVID-19 pandemic has driven regional governments to put more effort into integrating ICT into their operations¹³. Administrative processes and service delivery can be better streamlined with the adoption of ICT into national and regional governance systems. ICT can facilitate establishment of one-stop service windows when accessing government services, because if government data are integrated, transaction times should be reduced, leading to more productive government ministries and agencies, and the geographic location at which a transaction begins will not affect delivery of a service. Governments must employ a 'coherent use of digital technologies by promoting the use of compatible technologies and the proper update of

¹³ Making progress integrating ICT depends on economic capacity and the success of public-private partnerships.

ICT frameworks across policy areas and levels of government' (OECD, CAF and ECLAC 2018, 169). Adoption of ICT will support systemic transformational change.

Beyond the use of digital tools to deliver citizen-centric services, governments can no longer afford to separate efficiency from societal policy objectives. To that end, they need digital technologies to support policy design, implementation and evaluation. They must pursue these goals while developing and reinforcing capacities to manage and monitor digital strategies and assess their outcomes. This process goes beyond the deployment of technologies. It encompasses a technical knowledge in the acquisition of ICTs and well-structured governance with strong leadership. Further, it demands rethinking services to empower all citizens, not only those who are technologically savvy or connected (OECD, CAF and ECLAC 2018, 169)

Bringing It All Together

ACS stakeholders envisaged a role for evaluation and recognized the benefits of regional thinking and the importance of planning for sustainable development. From the United Nations 2030 Agenda, the SDGs that are focused on health, food security, transportation and supply chains were the priority for countries in the Greater Caribbean in the management phase of the COVID-19 pandemic. However, in the recovery phase, these countries would prioritize the SDGs focused on developing partnerships to leverage economies of scale, regarding preservation of the environment and economic development.

Transformation of the national and regional governance systems was deemed essential for the Greater Caribbean. The principles of the systems approach were used to identify four pathways to realize that transformation. First, a four-tier education drive using transformative change practices and transformative evaluation with government leadership, senior public servants, specialists and administrative staff was proposed. A multidisciplinary group of public servants from all levels of government would help build the evaluation culture and coach other personnel to contribute to and use the internal M&E systems in ministries and agencies. Leveraging established data collection and management systems would help integrate evaluation into all components of national governance systems.

The second proposed pathway was a reformed, inclusive, strategic planning process with advocacy for implementation of a whole-of-government approach and an emphasis on coordination and collaboration across

the national system. A meta-analysis of regional studies completed in the last 10 to 15 years was suggested to identify recommendations to improve governance systems. The third pathway was implementation of more inclusive policy development and M&E processes. The importance of alignment of national policies with regional policies to boost regional development and improve outcomes of partnerships internal and external to the region was emphasized. The fourth pathway was to increase use of technology to improve service delivery of governance systems and open channels of communication between governments and citizens. The pathways should yield citizen-centric services and increase citizen trust in the governments of the region while addressing the complexity of governance systems that would hinder the change process.

Building and accepting a culture of evidence is critical in an environment of shrinking resources to help determine the most efficient and effective use of resources. The highest level of government officials must be engaged in the transformation practices and demand and use evaluation findings.

Countries sharing the economic space of the Greater Caribbean must build an awareness of the institutions in the new international governance after COVID-19 to promote regional initiatives and defend the region's interests and aspirations, as it navigates renewed international cooperation. It is essential that the processes of regional integration be strengthened.

References

- Association of Caribbean States. 2017. About the Association of Caribbean States. <http://www.acs-aec.org/index.php?q=about-the-acs>
- Baptiste, L., and T. Moss. 2017. A British and Dutch Caribbean overseas territories training evaluation case study: An HIV/AIDS workforce training perspective. In: *Training Initiatives and Strategies for the Modern Workforce*, (eds.) S. Frasad and F. Prasuhn, 19–40. Hershey, PA: IGI Global.
- Baptiste, Lennise, Villiamu Iese, Valerie Gordon et al. 2019. The transformative agenda for evaluation in small island developing states: The Caribbean and the Pacific. In: *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder, 71–88. Exeter, UK: IDEAS.
- Bolinson, C., Donna Mertens and Engineers Without Borders Canada. 2019. *A transformative evaluation toolkit for the impact investing sector: How entrepreneurs and impact investors can deepen their data collection*. Toronto: Engineers Without Borders Canada.

- CARICOM (Caribbean Community). 2020. Address by CARICOM chair to inter-sessional summit of the Organisation of African, Caribbean and Pacific States (OACPS). <https://tinyurl.com/s6hh2my9>
- ECLAC (Economic Commission for Latin America and the Caribbean). 2018. *Latin American outlook 2018: Rethinking institutions for development*. Santiago: ECLAC.
- ECLAC (Economic Commission for Latin America and the Caribbean). 2020a. Borrowing is not an option for Caribbean countries, access to concessional funding and debt relief is urgently needed to face COVID-19 crisis. <https://tinyurl.com/wbzs9tz>
- ECLAC (Economic Commission for Latin America and the Caribbean). 2020b. *Building a new future: Transformative recovery with equality and sustainability* (LC/SES.38/3-P/Rev.1). Santiago: ECLAC.
- ECLAC and PAHO (Pan American Health Organisation). 2020. Controlling the pandemic requires convergence and coordination between health, economic, social and productive policies. <https://tinyurl.com/2ev9ujv5>
- Feinstein, Osvaldo. 2019. Dynamic evaluation for transformational change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvina Salinas Mulder, 71–88. Exeter, UK: IDEAS.
- Hargreaves, M. 2010. *Evaluating system change: A planning guide*. Princeton, NJ: Mathematica Policy Research.
- Levaggi, Ariel Gonzales. 2020. COVID-19 and geopolitics in Latin America and the Caribbean. <https://tinyurl.com/3z9wavtm>
- Magro, Christina, and Rob D. van den Berg. 2019. Systems evaluations for transformational change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvina Salinas Mulder, 71–88. Exeter, UK: IDEAS.
- Matera, Michael, and Maria de Lourdes Despradel. 2020. *Latin America and the Caribbean in the new decade: How did we get here?* Center for Strategic and International Studies.
- OECD (Organisation for Economic Co-operation and Development). 2017. *Production transformation policy reviews: Actions to succeed in a changing world*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/9789264276628-en>
- OECD, CAF (Development Bank of Latin America) and ECLAC. 2018. *Latin American economic outlook 2018. Rethinking institutions for development*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/leo-2018-en>
- Ofir, Zenda. 2018. Towards Transformation, 1. The buzzword. Evaluation for Transformation blog. 10 July. <https://tinyurl.com/y24hdp9v>
- Ofir, Zenda, Gerald Singh, Emile Beauchamp, Dorothy Lucks, Stefano D'Errico and Kassem El-Saddick. 2019. From monitoring goals to systems-informed evaluation: Insights from SDG14. *IIED Briefing March*. <https://pubs.iiied.org/pdfs/17706IIED.pdf>

- Persad, Armika. 2020. An examination of the effects of COVID-19 on the transportation sector in the greater Caribbean – the case of Panama & Jamaica. <https://tinyurl.com/65wxujyw>
- Polastro, Ricardo. 2014. Evaluating humanitarian action in real time. Recent practices, challenges and innovation. *Canadian Journal of Program Evaluation/ La Revue Canadienne d'évaluation de programme* 29 (1): 118–134. <https://tinyurl.com/y546sw8z>
- Taut, Sandra, and Dieter Brauns. 2003. Resistance to evaluation: A psychological perspective. *Evaluation* 9 (3): 247–264.
- United Nations. 2016. UN chief launches first report to track Sustainable Development Goals. <https://tinyurl.com/55embrw>
- Wiltshire, Rosina. 2015. Towards a Caribbean multi-country assessment: A base document for stakeholder and partner consultations.
- World Bank. 2020. Countries can take steps now to rebuild from COVID-19. <https://tinyurl.com/jf6kctcb>

PART III

**PROFESSIONAL-
IZATION**





CHAPTER 8

To Be or Not to Be an Evaluator for Transformational Change: Perspectives from the Global South

PABLO RODRÍGUEZ-BILELLA, SILVIA SALINAS MULDER
AND SONAL ZAVERI

Abstract. In the current global neoliberal context, evaluation runs the risk of becoming another service that gives answers wanted by those who pay for it. Being a transformative evaluator entails extending the focus of action to contribute to public good, broadening its interest towards medium- and long-term results, and investigating the root causes of those social problems that programmes and policies aim to deal with. This chapter introduces a theoretical framework on transformative evaluation based on theory and practice from the Global South. For that, it discusses a competencies profile for gender- transformative, context-relevant evaluations, a comprehensive approach built in Latin America. Then, selected cases are presented to identify the factors and evaluator competencies that facilitate usable evaluation and evaluations aimed at social betterment. The last section discusses the complexities underlying frequently invisible power issues and relations and the need to fine-tune one's ability to identify and address them in evaluations. The chapter stresses the importance of redefining the role and competencies needed to enhance the transformative potential of evaluators, ensuring gender responsiveness and power awareness under culturally diverse and complex realities, identifying evidence-based strategies and actions to conduct evaluations that have a positive impact on people's lives.

Introduction: Changing Evaluation Paradigms

Evaluators usually do not go into detail on how their work can improve people's lives. They assume that their responsibility does not extend beyond selecting the appropriate methodology or method capable of influencing decision-making. In the global neoliberal context, that behaviour increases the risk of making evaluation another service that answers the questions of those who pay. Although inclusive, participatory evaluations are gaining ground, many evaluations concentrate excessively on efficiency, effectiveness and measurable results on a short-term basis rather than contributing to democratic, transformative and participatory purposes that the evaluation community holds as central.

Being a transformative evaluator entails extending the focus of action of the evaluation to contribute to public good; broadening it towards medium- and long-term results and to unexpected consequences of development interventions and investigating the causes of some social problems that programmes, projects and policies are designed to address. For this, competencies are required that go beyond analysing performance aspects – those that allow the transformations that reduce poverty and inequalities among the most disadvantaged groups in society to be identified, addressed and facilitated. It is also necessary to develop competencies to 'learn to and lead for change' in contexts in which it is imperative to induce cultural changes to transform unequal power relationships and perverse social norms.

Some premises that can help evaluations become learning and transformation processes and help their results redefine strategies for greater transformational impact are that:

- evaluations do not take place in a vacuum. There are political, economic, cultural and even technological forces that can facilitate or inhibit results. Evaluators must be aware of their existence to anticipate possible evaluation scenarios and develop strategies that challenge the status quo and change paradigms.
- evaluation is a political activity, not a process devoid of value and interest. Evaluation processes are part of the 'change we want to see,' and evaluators help achieve it.
- challenging power relations begins by breaking the hierarchy in the relationship between the evaluator and the 'evaluated' entity.
- evaluations are inclusive, mutually educational, empowering processes.

- methodological credibility is essential to support the findings. New techniques and tools should be explored while preserving the rigour of the analyses at the same time that conventional understanding of rigour is redefined.
- transformation is context specific.

Approaches such as the Blue Marble Evaluation¹ also bring different perspectives that seek to look beyond the contexts of projects and programmes, beyond national borders and between silos and sectors ‘to connect the global with the local, connect the human and ecological, and connect evaluative thinking and methods with those trying to bring about global systems transformation’ (Patton 2020, 1).

The Decade of Action towards the Sustainable Development Goals (SDGs) and the worldwide disruptions that have resulted from the COVID-19 pandemic have highlighted the relevance and importance of evaluation in contributing to transformative purposes such as solving deep structural problems, challenging inequalities, overcoming barriers that inhibit agility instead of moving quickly towards achievement of the SDGs, providing evidence that supports scaling development models and boosting social innovation. Are we evaluators ready to take the challenge? In these times of volatility, uncertainty, complexity and ambiguity, the challenge is even greater, but we definitely cannot continue working under traditional professional paradigms. We need new competencies to understand complex realities, to be people centred, ethically accountable, transformative and resilient. We need to become agents of evaluation with a purpose.

This chapter will help the reader distinguish transformative evaluation from conventional approaches, as well as specify further desirable correlations that make transformative evaluation a robust and relevant approach, addressing subtle differences from other change-oriented approaches. After making that fundamental presentation, we introduce a competencies profile for evaluators who work in diverse contexts, address gender transformation and challenge equity-related power imbalances. This comprehensive approach, developed in Latin America and shared with South Asian and African evaluators, integrates the technical, ethical and political dimensions. The third section presents selected cases from *Leaving a Footprint: Stories of Evaluations That Made a Difference* (Rodríguez-Bilella

¹ Blue Marble Evaluation focuses on transforming evaluation to evaluate the transformations necessary to reverse damage from climate change and make human life on Earth more sustainable and equitable (Patton 2020).

and Tapella 2018) that provide evidence-based guidance on factors and evaluator competencies that facilitate usable evaluation and contribute to the body of knowledge of evaluations aimed at social betterment. The final section describes the complexities underlying frequently invisible questions of power and power relationships and the need to fine-tune one's ability to identify and address them in evaluations conducive to the occurrence of transformations.

It is hoped that this chapter will raise awareness of the importance of redefining the roles and competencies needed to enhance the transformative potential of evaluators, ensure inclusiveness (exploring mainly gender issues, because of the authors' previous work) and power awareness under culturally diverse and complex realities and identify evidence-based strategies and actions that can be used to conduct evaluations that improve people's lives.

From Conventional to Transformative: Evaluation with a Purpose

Evaluation is described as transformational herein, to contrast it with how evaluation is conventionally conducted. Conventional evaluation is driven primarily by a positivist world view that emphasizes observation and reason – to assess processes and outcomes. The definitions of evaluation describe it as a neutral exercise. Scriven's (1991) definition of judging merit, worth, value or significance is useful to distinguish evaluation from research, but many evaluators interpret this definition as saying that merit and values are universal and similar and that evaluators will be 'objective' in their assessments. Rossi's (2004) definition, which is also popular, describes use of social research methods to systematically investigate the effectiveness of social interventions. The Organisation for Economic Co-operation and Development Development Assistance Committee (n.d.) criteria also emphasize the systematic and objective assessment of an ongoing or completed project, programme or policy and its design, implementation and results.

One may argue that these definitions represent a pragmatic rather than a 'pure' positivist approach, but they are all based on the premise that data should inform what works or does not and in doing so establish a causal relationship. To do so, the most appropriate tools of science and technology for these tasks are applied. What is important for this discussion is the overriding evaluative intention of accountability in conventional evaluations. There

may be less understanding of or attention paid to assumptions underlying programme design, how implementation changes in unpredictable ways and how diverse populations perceive and receive results. To illustrate, a transformative approach to evaluating a cash transfer programme would want a deep understanding of *which change mattered to whom, why and whether that was 'enough'*, acknowledging structural inequities in the lives of poor women. Conventional tools, mostly Western driven, may not assess these living realities. Transformative evaluations question whether 'verifiable' results had real impacts on the power structures that dominate the lives of the marginalized and vulnerable.

The post-positivist approach questioned the conventional framing of evaluation and urged consideration of alternate views to make judgments in evaluation. Evaluators began questioning the lack of acknowledgment for the context, the evaluator's own biases, the lack of emphasis on the voices of those affected by interventions and the complexity of interactions that need to be addressed. Participatory, developmental, systems and goal-free evaluation approaches that are not positivist have gained in popularity because they address the multiple, ever-changing realities of life. Alternative definitions began to emerge, one from the Global South, stating that to evaluate is to assess the overall impact of a *social change* intervention against an explicit set of goals and objectives and to determine what works and what does not (Batliwala and Pittman 2010).

Transformative evaluation is an approach that has been in development for longer than 15 years (Cooper 2013), which can be understood as an expression in the evaluation field of the need for transformation in society. Freire's (1994) call for equality mobilized communities more than four decades ago, and work by participatory and feminist evaluators (Kabeer 2001) has championed the need to include voices that are often marginalized. The demand for transformation is also a product of our complex times, as we grapple with persistent problems of poverty, gender inequities and discrimination, as well as emerging and urgent phenomena that affect societies globally, such as COVID-19. The 2030 Agenda's bold recommendation that 'no one be left behind' spotlights inequities and inevitably demands that interventions be transformative in their approach.

The 2030 Agenda, unlike the Millennium Development Goals, uses a transformative lens, urging a 'people-centered, human rights and gender equity' approach with a particular focus on the poorest and most vulnerable and ensuring that no one is left behind. To do so, it posits that change must be transformative, which means attacking the root causes of discrimination because they generate and reproduce economic, social, political and

environmental problems and inequities (UNRISD 2015). In other words, the demand is to address root causes, not just symptoms of the problem.

Conventional and transformative evaluation represent different world views. A world view is composed of our beliefs, values and assumptions about the world we live in and interact with. In other words, we use our world view to make sense of our experiences in the world we live in. Practically, in evaluation, this means that our world view guides *our decisions* about the definition of any object of study, what questions to ask, how to ask them and how to interpret our findings. World views work in the background, and we become acutely aware of them usually when confronted with an alternative world view. The conventional methodologies are grounded in philosophical assumptions commonly known as ‘positivist’, whereas the transformative methodologies represent a more systemic and interpretive point of view. Positivist and more conventional evaluators are grounded in Newtonian assumptions that favour predictability, replicability and the observable and seek mostly linear, instructive causal connections. Transformative evaluators, on the other hand, value *what* phenomena are observed; suggest a more nuanced understanding of these observations; view the context in which studies are conducted as unique (often not replicable) and look for narratives, correlations and explanations rather than causal connections. This complex view of the world brings uncertainty and recommends constant adaptation; this is the uncertainty of different realities that we must acknowledge and address.

The assumptions of these two world views are different in terms of what counts as real (ontology), how we know and make sense of our what we know (epistemology) and values and beliefs (axiology), so when we talk of conventional or transformative evaluation, we need to be aware that these are different world views that, in turn, influence how we evaluate.

A transformative lens acknowledges that there are multiple perspectives, each expressing a different reality, where some voices are heard, and some are not. It is this inequality of whose voice matters that underpins our discussion about transformative evaluation. The underlying principle of social justice (Mertens 2007) demands that change processes be assessed in terms of inclusion, equity, sustainability and fairness.

Such thinking should encourage evaluators to question the purpose of evaluation, why we do what we do, who it serves and who it benefits. Evaluators are accustomed to discussions about approaches, methodology and tools but seldom about evaluation’s contribution (or not) to equitable, just societal change. Should evaluations confine themselves to assessing outcomes and impact, or should evaluation be an empowering exercise that

addresses the needs of all, particularly those who are being left behind? This is an important shift from conventional evaluation, whose purposes have traditionally been for accountability and learning and to increase an agency's capability (Chelimsky 2006). The emphasis has been on accountability, usually upward, coining terms such as 'value for money' and 'social return on investment'. Even when downward accountability is acknowledged, participation of those that a programme or policy most affects can be tokenistic, with predetermined theories of change based on assumptions that do not address the complexity of power asymmetries and social change processes.

The purpose of transformative evaluation is learning, and accountability is redefined as the democratic sharing of responsibility. We move away from narrowly defined technocratic uses of evaluation to how evaluation benefits the lives of the people most affected. It requires that people (rather than 'beneficiaries') be involved in the evaluation and learning process and not be considered merely as sources from which to gather data. For instance, people that programmes affect actively and purposefully contribute to what data needs to be collected, engage in discussions about why they need certain data and discuss how the data collected and analysed will be used to make decisions. The role of the transformative evaluator is facilitative rather than directive – being in charge of the evaluation process and making judgements as an expert evaluator.

Another characteristic of transformative evaluation is how it understands the nature of change. Transformative thinking is a paradigm shift – understanding that change itself is complex and long term; that we may need to have a more nuanced world view of how change occurs and that the methodology and theoretical framework required in practice is radical, unconventional and innovative. Changes in outcomes are not just incremental or even reform-based, but truly transformative, addressing the root causes of power inequities. The types of questions we ask as evaluators, the purpose of doing so, addressing power (or not), and our actions and tools will depend upon the type of change we evaluate. Incremental change is essentially about improving performance, such as evaluating the expansion of an existing immunization programme (following the same protocols), or 'within the box' change. Reform change is change 'outside the box', where new rules are addressed, usually with policy reform or some sort of restructuring. This could refer to work conditions for female workers in an industry that has previously not addressed workers' rights or the different needs of female workers – superficial changes that improve but do not address root problems. Transformative change refers to a fundamentally new way of addressing the phenomena and could be an innovation or an experiment.

It 'questions the box' itself. An example of such transformative change is ending apartheid. The following example may explain the different types of changes: incremental change would be making available sources of energy more efficient, reform-related change would be advocating for solar energy and excluding fossil-based fuels and transformative change would be changing our lifestyle dramatically so that we live an energy-frugal lifestyle or negotiating energy consumption with Indigenous people whose land provides us the energy resources (table 8.1).

Transformative change redefines accountability as much for the people with whom the project works as for those who are left out. Those who adhere to a transformational stance are likely to seek longer funding cycles and have the patience to work through small transformational changes, such as women speaking up in meetings or, better still, a poor woman who rarely comes to a village meeting speaking up.

To summarize, the transformative evaluator has a fundamentally different understanding of what development, participation, empowerment and

Table 8.1 Types of Change

Change type	Incremental	Reform	Transformation
Core question	How can we do more of the same? Are we doing things right?	What rules should we create? What structures and processes do we need?	How do I make sense of this? What is the purpose? How do we know what is best?
Purpose	Improve performance	Understand and change the system and its parts	Innovate and create previously unimagined possibilities
Power and relationships	Confirm existing rules	Open rules to revision	Open issue to creation of new ways of thinking about power
Action logic	Project implementation	Piloting	Innovating
Archetypal actions	Copying, duplicating, mimicking	Changing policy, adjusting, adapting	Visioning, experimenting, inventing
Tools	Negotiation	Mediation	Envisioning

Source: Adapted from 2017 SDG Transformations Forum.

accountability mean. This in turn determines what we measure, how we do so and who we include in participation. There is a deep belief in ownership and sustainability. A transformative evaluator will always ask what we mean by impact, who we impact and who we leave behind. Specifically, did it make any meaningful difference, was there social justice?

Evaluators as Change Actors: A Competency-Based, Gender Transformative Approach

Evaluations are conventionally expected to be neutral and power blind, but the conscious or unconscious biases of the evaluator implicitly or explicitly frame evaluations. Evaluation in its present form has had a long history in the Global North, centred mostly on white men's contributions. The myths and assumptions that govern conventional development evaluation and are male-biased and rooted in misconceptions about the neutrality of social conditions for development are exacerbated in contexts in which political interests; patriarchal conceptions and values and ethnic, religious and other fundamentalisms influence evaluation decisions. Likewise, there is a widespread conception that gender issues are reduced to programmes and projects for women. Although evolving towards a more gender-fair approach, the idea prevails that major development problems, such as climate change, food security, malnutrition and infrastructure construction have nothing to do with these inequalities. Transformative evaluation introduces gender as a quality criterion of evaluations; boosts the importance of ethics and accountability; emphasizes people-centred evaluation practice; pays particular attention to people who experience any form of inequality, discrimination or vulnerability; and recognizes evaluators as change actors, considering their opportunity to empower actors in the evaluation process; truly assess the transformative nature of policies, programmes or projects and influence evaluation design, process, analysis and use.

Gender analysis and feminist theory have made fundamental contributions to the practice of transformative evaluation in terms of applying the principles of human rights, equality, participation and non-discrimination that allow rigorous analysis to transform the roots of gender inequality. Despite the formal advances, in practice, the evaluation standards do not emphasize the importance of examining and challenging unequal gender power relations or determine the obligation to do so in a context in which there are still false dichotomies between subjectivity and objectivity, qualitative and quantitative, effectiveness and efficiency, spending and

investment, and North and South that limit the measurement of results in terms of social change. Gender blindness in evaluations neglects the fact that reducing gender gaps is not only a matter of justice, but also a key factor to boost development, productivity and poverty reduction and thus a key quality criterion of evaluations.

Motivated by the idea to 'make gender a quality criterion of evaluations', the Network of Latin American and Caribbean Women in Management and Centers for Learning on Evaluation and Results Latin America and the Caribbean implemented a virtual course titled 'From Conventions to Innovations: Agents of Change to Promote a Gender Approach in Evaluations' in 2015. This was one of the winning initiatives of the global Innovation Challenge: A Focus on Equity and Gender Responsiveness in Evaluations that the EvalPartners alliance launched. Its innovative approach included a participatory preparatory process with a multidisciplinary group of Colombian stakeholders that, using the Development of a Curriculum methodology (Norton 1997), drafted a first set of evaluation competencies for different actors related to teaching, contracting, implementing and using evaluation. The course, which brought together a powerful, committed group of professionals from government, academia and civil society, as well as evaluation practitioners, ended with an initial collectively developed profile of competencies to promote evaluation with a gender approach, articulating the technical, political and ethical dimensions (Amariles, Salinas and Grandjean 2016).

In November 2015, under the inspiration of the launch of EvalGender+ in Kathmandu, Nepal, and in the context of the promulgation of the SDGs and their challenging mandate to leave no one behind, a collaborative initiative that concluded with a Decalogue of Evaluation with a Gender Perspective emerged from among several Spanish-speaking feminist evaluators (figure 8.1) (Salinas Mulder 2015). The Decalogue constitutes a frame of reference from which to promote the gender approach as a quality criterion for evaluations and makes it easier to follow the discussion on the competencies needed to evaluate with a gender perspective from a transformative, culturally relevant perspective, including competencies necessary to influence construction of an enabling environment that demands and advocates for institutionalization of the gender approach in evaluations of development (Amariles et al. 2015).

From January to August 2017, the Development of a Culturally Relevant Curriculum on Transformative Gender Evaluation project was implemented under the EvalPartners Peer-to-Peer initiative, in which four regional networks of evaluation of the South participated (Latin American Network for

Figure 8.1 Decalogue of Evaluation with a Gender Perspective

This Decalogue seeks to help avoid the evaporation and technocratization of gender issues in evaluation in order to contribute to more gender-transformative practices.

- 1 Recognizes and values the political dimension of evaluation to contribute to transform gender inequalities and promote social justice.
- 2 Assumes that public policies, programmes, projects and their evaluation are not gender neutral.
- 3 Acknowledges evaluation as a quality criterion, noting that a gender perspective should be applied to all kinds of policies, programmes and projects.
- 4 Implies questioning gender power relations and analysing results and processes.
- 5 Proposes a holistic approach: one that looks at people, organizations/ institutions and their environments.
- 6 Promotes participatory and collaborative work to build collective knowledge and empowerment.
- 7 Focuses on accountability, learning, improvement and advocacy with a view to transforming gender inequalities.
- 8 Generates analysis, conclusions, recommendations and lessons learned to promote changes in gender relations.
- 9 Adopts and adapts gender analysis and other tools and methodologies to local contexts, languages and the cultural characteristics of communities.
- 10 Analyses how gender inequality intersects with other inequalities.

Source: Adapted from Amariles, Salinas, Espinosa et al. (2015).

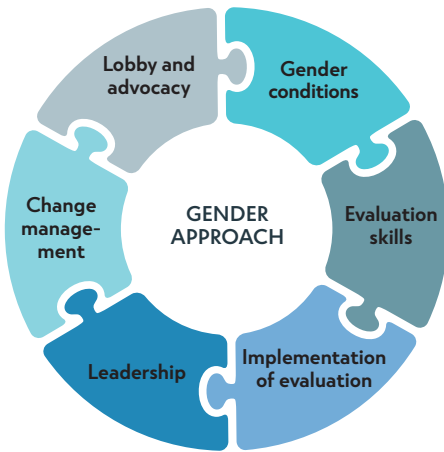
Systematization, Monitoring and Evaluation; Network of Latin American and Caribbean Women in Management; Africa Gender and Development Evaluators Network; Community of Evaluators, South Asia). The objective was to share experiences, considering the realities of the regions, and incorporate an approach of cultural relevance into the development of evaluation competencies.

In the first stage of the project, the Latin American team developed a widely participatory process through which two main results were achieved: (1) administration of a survey to determine whether there are any capacity development programmes on gender-transformative evaluation in Latin America and the Caribbean (Spanish and Portuguese speaking) and perceptions of the evaluation community on this topic and (2) participatory refinement of the Integral Profile of Competencies for Evaluators from a

Gender Transformative Approach with Cultural Relevance that starts in Latin America and the Caribbean and is applicable to other regions (Amariles, Salinas Mulder and Rodríguez-Bilella 2018).

The competency-based approach that was developed with diverse participation between 2015 and 2017 through different initiatives discards the previously prevailing assumption that it was sufficient to know to be able to do, highlighting the importance of skills and attitudes to transform knowledge into action and, at the same time, generate the conditions of viability, ethical relevance and coherence of the evaluation process. Thus, the competency profile addresses performance and the real capacity to achieve an objective, solve a problem or achieve a result in a specific context. It

Figure 8.2 Competencies for Evaluators from a Gender-Transformative Approach with Cultural Relevance



combines 'ways of knowing', 'ways of doing' and 'ways of being', developing a new paradigm for evaluation practitioners from a holistic, comprehensive approach.

The Integral Profile of Competencies for Evaluators from a Gender-Transformative Approach with Cultural Relevance proposal is organized as a system with seven dimensions designed to articulate a comprehensive approach not only from the themes or contents identified as necessary, but also from an integrated approach that addresses the political, ethical and technical aspects of evaluations (figure 8.2).

- *General conditions* refer to the cross-cutting aspects that build on the positioning of the evaluator, such as their critical knowledge of the context, conceptions of otherness and reflective capacity – in the words of one of the expert reviewers, 'the hidden profile of an evaluator'.
- *Evaluation skills* not only reflect traditional approaches to evaluation competencies (the knowledge and ability to analyse the current regulatory, institutional and policy framework related to evaluation), but also include new 'technical' competencies such as systems perspectives and adaptive approaches that are linked to a transformational approach.

- *Implementation of the evaluation* refers to the realization of the evaluation and includes the whole cycle, from planning to communication of results and recommendations.
- *Gender perspective or approach* is the heart of the profile proposal, a competency that must be integrated into all other competencies for its effective implementation. It includes a wide range of knowledge, skills and attitudes that the evaluators must develop.
- *Leadership* assumes that the environment is frequently not receptive or favourable to including a gender perspective in evaluations. Thus, evaluators committed to including a gender perspective in their practice must also play a leadership and change-actor role by promoting recognition and operationalization of gender as a quality criterion for development evaluations.
- *Change management* is one of the most important contributions and refers particularly to one of the main challenges that evaluation faces today: its use. This dimension goes much further than a necessary follow-up; it identifies the knowledge, attitudes and skills required to motivate and support implementation of changes based on the evaluation findings and recommendations.
- *Lobbying and advocacy*: the key component of the proposal has to do with connecting evaluation with the possibility of advocating for and influencing changes based on the evaluation findings and recommendations. Adoption of a system perspective is a critical component for producing substantive change in various areas and levels, from public policies to gender-blind ways of and criteria for traditional evaluation.

This profile is a work in progress. The re-emergence of conservative and even fundamentalist forces worldwide, COVID-19 and its general and gender-related consequences, expanding inequality gaps and intersected oppressions, increasing multifaceted violence, climate change catastrophes and migrations are among the complex contemporary realities that we evaluators must fully understand and address. New and dynamic evaluator competency profiles must better equip practitioners to play a transformative role.

Leaving a Footprint: Inspiration from Evaluators Who Made a Difference

The commitment to building a new set of competencies for evaluation must meet users' needs, as well as evaluation quality standards based on credible evidence. In bridging the gap between theory and practice, or between resources invested in evaluation and its use, it may be useful to reflect on evaluators' competencies based on stories of evaluations that have made a difference². Analysing evaluation stories can help identify factors that facilitate development of useful evaluations and contribute to the body of knowledge of evaluations aimed at social betterment (Mark, Henry and Julnes 2000), that is, evaluations that improve people's lives.

Evaluation can transform the lives of those whom programmes and policies affect *by providing a space for their voices and their expression*, contributing to their inclusion in decision-makers' mental models. This situation is enhanced in the many cases in which decision makers do not have a close connection to the many realities of programme participants, not knowing their needs or contexts. The story of the qualitative evaluation of the Progres/Oportunidades (Mexico) programme illustrates how evaluation identified language barriers that prevented very poor natives from benefiting from a money transfer programme. Changes to the programme allowed communication in local languages, which greatly increased the ability of people to understand the programme's requirements (e.g. children's regular school attendance) and therefore to benefit from the money transfer the programme offered.

Programmes and development policies designed to improve people's lives are increasingly being expected to be based on credible evidence. *A key competency for evaluators is their capacity to choose the best way to generate believable, convincing information*, given that what is 'believable' depends on the situation and the specific actors. Evaluation credibility may be achieved in different ways – sometimes by using an approach that helps the process to be perceived as methodologically rigorous, other times focusing on and understanding the perspectives of the most relevant actors in the intervention and other times through active participation of users in the evaluation process.

An example of active participation of users in the evaluation process was the participatory evaluation experience in the cancer prevention and

² This section is based on Rodríguez-Bilella and Tapella (2018) and Perrin et al. (2015).

care programme in Valle de la Estrella in Costa Rica, where regional technical teams were involved and deeply interested in understanding how the evaluated programme worked in their area. In contrast, higher authorities limited their participation to approving the evaluation. In this way, recommendations at regional and local levels were applied soon after the evaluation finished, whereas general recommendations – dependent upon higher authorities – have not yet been applied.

In every evaluation that makes a difference, *the technical ability, rigour and competence of the evaluator or evaluation team is highly significant*. In the evaluation of the Mexican programme, the key factor was the evaluators' anthropological approach, whereas in the evaluation of the cancer prevention programme in Costa Rica, the interdisciplinary nature of the participatory evaluation was very important. Beyond the technical rigour, *communication of the evaluation results to relevant actors is becoming increasingly important*. Communication draws attention to the type of report used, adapting language to different audiences and generating lessons learned that fall within the ability of the organization's ability to respond.

Evaluators often try to keep a certain distance from evaluated programmes to protect their independence, but this increases the likelihood that the evaluation becomes distant and irrelevant for those who need to act on the results. *Being close to the evaluated programmes and their actors gives evaluators opportunities to make a difference through the evaluation process*. This recognizes that the benefits and impacts of evaluation emerge as much as – or even more than – from how an evaluation is conducted (usefulness of the process) as in relation to its findings (usefulness of its results) (Cousins, Whitmore and Shulha 2013).

Including and involving users and participants in collection and use of evaluation data is a powerful way to gain a better understanding of those data. *A powerful and desired competency is to lead participants to take responsibility for the evaluation and for the change and transformation that follows*. Active participation in the evaluation process helps develop better understanding of evaluation and contributes to commitment and use. As the Costa Rica case shows, *the more participatory the evaluation is, the more necessary it is to ensure the willingness and motivation of the most relevant actors in the intervention* (participants, local technicians, officers) in order to promote the impact of the evaluation and for it to make a difference.

In the early stages of the evaluation, it is common for most of the intervention participants, as well as the actors who implement the programmes

(e.g. field technicians, officers in charge of the implementation), to consider evaluation from a point of view of control and accountability. Generally, the start of an evaluation process does not create excitement or expectations connected to the learning dimension. The situation changes *when the evaluator or evaluation team is able to show through their words and actions that evaluation has the potential to improve programmes*, overcoming narrow views connected with monitoring and control, accountability, rewards and sanctions. The evaluation story in Mexico illustrates how indigenous women were invited to participate in the evaluation with the intention of decreasing anxiety, without explicitly mentioning that they were being involved in an evaluation process, which was made clear soon after the women arrived.

The impact of an evaluation can be increased, as a much-desired competency of an evaluator, by having champions who can influence those who make key decisions and necessary changes. It is common that the people who have real authority to make decisions are external to the programme and have not participated in the evaluation process. Thus, even though the programme staff and the directors who took part in the evaluation are committed to improving the programme, other interested parties need to be convinced that the changes are necessary. Champions in evaluation usually are people who care deeply for the affected families and communities and also have an influence on others who are able to make decisions, playing a fundamental role so that the changes can take place. In the example of the evaluation of the Progresia/Oportunidades programme in Mexico, an actor who believed in the potential of the evaluation effort and facilitated implementation of some of the suggested recommendations played that role.

The idea of speaking truth to power may be naïve and insufficient if the inherent political nature of evaluation is not recognized. This entails extending the focus of action of the evaluation to contributing to the public good, broadening its interest towards medium- and long-term results (including unexpected consequences of development interventions) and investigating the causes of some social problems that programmes and policies are designed to address. Giving evidence to subjects of the political intervention entails 'addressing the truth to the powerless', which may be considered a new and relevant competency for evaluators. This requires considering them as legitimate stakeholders in the evaluation results and empowering them to speak for themselves and act on their own benefit. Developing strategies for that is a challenge, as well as a rich field of development for evaluators.

This section discussed seven principles for credible evaluation: being inclusive, selecting good data, using rigorous techniques, communicating

results, being close to the evaluand, having a champion and focusing on use. Although some of these competencies have been around for a while, their implementation and inclusion in evaluation competency profiles is not always harmonized and consolidated, especially in the Global South.

The Devil Is in the Details: Final Reflections

Conventional evaluation tends to produce unequal power relationships and reinforces established unequal power relationships. Power is a tricky and invisible issue that 'hides' in how we relate, view, think and analyse. In this final section, we briefly address some key power dynamics in evaluation that indicate the need to expand evaluator competencies to be better prepared to identify and disentangle power 'knots'.

- *Reflexivity*: This is not a frequently considered competency. Reflexivity and self-awareness help us recognize that we see and evaluate through the eyes of our own history and environment and that our view of reality is always partial and slanted. We represent power structures and relationships that can be expressed in subtle manners. Even gender expertise does not necessarily challenge machismo and patriarchal relationships; addressing power relationships requires changing (one's) culture.
- *Ethics*: Despite good intentions, ethics frequently focuses on formal compliance, with little attention paid to accountability mechanisms that capture and analyse how ethical topics and even dilemmas are tackled in the field. The lack of contextualization and understanding of local dynamics and power relationships worsens the unreflective compliance of formal ethical procedures. Unequal power relationships are frequently reproduced during evaluation processes, and even human rights can be violated by action or by omission during fieldwork. Beware of 'unethical ethics' and be transparent about the ethical dilemmas faced during the evaluation and decisions made under those circumstances.
- *Indicators*: There is an obsession with the idea of and need for success that several factors reinforced. This threatens reflexivity, accountability and learning and can cause a lack of reflexivity about whether 'success' is linked to transformations and improvements in people's lives. Even worse, complying with the established traditional indicators and targets can imply ignoring the policy's or

programme's underlying purpose and not observing participants' (women's) human rights. Success indicators may be misleading, and 'not everything that glitters is gold'; a new set of indicators must be identified for evaluations to be truly transformative.

- *Participation*: To ensure sufficient participation, sometimes not enough attention is paid to ethical premises and basic values that are perceived as potential inhibitors. Participation may entail different assumptions and approaches that do not automatically address power imbalances. Participation will be transformative if it is empowering, if people are treated not merely as 'key informants' but their presence is recognized, valued and somehow redistributed. Participation should be based on context awareness, intercultural dialogue, affirmative actions and trust-based conversations.

Becoming a transformative evaluator is not a one-time effort but a lifetime commitment that is driven essentially by two forces: the external factors that influence the diverse, complex and dynamic realities where evaluators perform and the reflexivity and self-awareness that guide us along the personal and professional path of permanent learning, growing and reinventing.

Acknowledgements

Authors are presented in alphabetical order because their contributions to the article were of equivalent value. The authors would like to acknowledge the contribution of Fabiola Amariles, who, in the IDEAS Global Assembly 2019 in Prague, had a leading role coordinating the workshop that inspired this joint article. We regret that she was not able to participate in this article due to personal reasons, but her ideas and inspiration accompanied us.

References

- Amariles, Fabiola, Silvia Salinas, Julia Espinosa et al. 2015. Decalogue of evaluation from a gender perspective. EvalGender+. <https://tinyurl.com/tc2pss8e>
- Amariles, Fabiola, Silvia Salinas and Solange Grandjean. 2016. Gender and equity in evaluations 2.0: From conventions to innovations. REDWIM/EvalPartners Project Innovation Challenge. <https://slideplayer.com/slide/8694953/>
- Amariles, Fabiola, Silvia Salinas Mulder and Pablo Rodríguez-Bilella. 2018. To 'leave no one behind': Towards capacity building in gender transformative evaluation with cultural relevance in Latin America.

- Batliwala, S., and A. Pittman. 2010. *Capturing change in women's realities: A critical overview of current monitoring and evaluation frameworks*. Toronto, Mexico City and Capetown: Association for Women's Rights in Development.
- Chelimsky, Eleanor. 2006. The purpose of evaluation in a democratic society. In: *The SAGE handbook of evaluation*, (eds.) I.F. Shaw, J.C. Green and M. M. Mark, 33–55. London: Sage.
- Cooper, Susan. 2013. Transformative evaluation: Organisational learning through participative practice. *The Learning Organization* 21 (2): 146–157.
- Cousins, J. Bradley, Elizabeth Whitmore and Lyn M. Shulha. 2013. Arguments for a common set of principles for collaborative inquiry in evaluation. *American Journal of Evaluation* 34 (1): 7–22.
- Freire, Paulo. 1994. *Pedagogy of hope*. New York: Continuum.
- OECD DAC (Organisation for Economic Co-operation and Development Development Assistance Committee). n.d. Evaluation criteria. <https://tinyurl.com/5kndt83r>
- Kabeer, Naila. 2001. *Discussing women's empowerment – theory and practice*. SIDA Studies no. 3. Stockholm: Sida.
- Mark, Melvin M., Gary T. Henry and George Julnes. 2000. *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.
- Mathison, Sandra. 2018. Does evaluation contribute to the public good? *Evaluation* 24 (1): 113–119.
- Mertens, Donna M. 2007. Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research* 1 (3): 212–225.
- Norton, Robert E. 1997. *DACUM Handbook*. Columbus: Center on Education and Training for Employment, Ohio State University. (Leadership Training Series No. 67).
- Patton, M. Q. 2020. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Perrin, Burt, Rochelle Zorzi, Pablo Rodríguez-Bilella et al. 2015. Evaluations that make a difference: Stories from around the world. EvalPartners. <https://tinyurl.com/392eytzz>
- Rodríguez-Bilella, Pablo, and Esteban Tapella. 2018. *Leaving a footprint: Evaluation stories that made a difference*. San Juan: Editorial UNSJ.
- Rossi, Peter. H., Mark W. Lipsey and Howard E. Freeman. 2004. *Evaluation: A systematic approach*, 7th ed. Thousand Oaks, CA: SAGE.
- Salinas Mulder, Silvia. 2015. Gender: A quality challenge for the 2030 Evaluation Agenda. <https://tinyurl.com/zcrx2muk>
- Scriven, Michael. 1991. *Evaluation thesaurus*, 4th ed. Thousand Oaks, CA: SAGE.
- SDG Transformations Forum. 2017. Transformation. <https://transformationsforum.net/transformation/>
- UNRISD (United Nations Research Institute for Social Development). 2015. 2030 Agenda for Sustainable Development. <https://tinyurl.com/4bdttufs>



CHAPTER 9

Avenues for Young and Emerging Evaluators' Journeys to 2030

KENZA BENNANI, MARIE-HÉLÈNE ADRIEN AND GERARDO SÁNCHEZ-ROMERO

Abstract. The Decade of Action offers young and emerging evaluators (YEEs) an opportunity to embark on transformative professional journeys to 2030. In this chapter, we share concrete avenues to help them conceptualize their career trajectories. We rely on the concept of 'professional identity work' to present a framework that defines and differentiates the various types of evaluator identities that YEEs could explore – through formal employment engagement or involvement with voluntary organizations for professional evaluation (VOPEs) and YEE networks. We posit that, in the coming decade, VOPEs and YEE networks could be considered as 'identity workspaces' that support YEEs in discovering, understanding and shaping who they are and can become in the era of the Sustainable Development Goals – as transformational evaluators in the making.

Heroes set out on a journey, transform into new and improved versions of themselves, and in the process encourage us all to follow in their footsteps.

–Allison and Goethals, *The Hero's Transformation* (2017)

Introduction

In his classic volume *The Hero with a Thousand Faces*, American mythographer Joseph Campbell (1949) outlined a set of principles for what he called 'The Hero's Journey'. Following Campbell, every hero myth is basically the same story, retold endlessly, in infinite variation. At heart, the story is always a *journey*, which implies leaving comfortable and ordinary surroundings, venturing into challenging and unfamiliar settings and returning *transformed* and empowered to advance society. This transformative journey usually starts with a *call to adventure* that presents the quest that the protagonist of the story must undertake, thereby establishing the stake of the game.

In this chapter, we issue such a call to adventure to the young and emerging evaluators (YEEs) of the world. We invite them to consider departing from their current professional configurations to embark on transformative professional journeys that will turn them into major protagonists of the Decade of Action that the United Nations has launched and place them front and centre in the 10 years left to deliver on the Sustainable Development Goals (SDGs).

Accelerating progress for achievement of the 2030 Agenda (United Nations General Assembly 2015) requires challenging the traditions of evaluation and rethinking our theory, practice and organization as a global community of evaluators so that these traditions can be better tailored to help realize the sustainable development ideals of our times (SDG Transformations Forum 2017). Adopted as an outcome of the IDEAS 2019 Global Assembly, the *Prague Declaration on Evaluation for Transformational Change* established the path to departing from the status quo by laying the groundwork for the realization of an ambitious transformational evaluation agenda. The declaration called for the development of new approaches in evaluation and recognized the potential contributions that 'new evaluators and collaborators from different disciplines and fields of work, including *young and emerging evaluators*' could make (IDEAS 2019, emphasis ours).

The Prague Assembly also saw the launch of *Evaluation for Transformational Change – Opportunities and Challenges for the Sustainable Development Goals*, the fifth book in the publication series of IDEAS

(Van den Berg, Magro and Salinas Mulder 2019), in which some of the founders of the EvalYouth Global Network presented a somewhat provocative essay on what they called 'a new evaluation revolution fuelled by youth' (Montrosse-Moorhead et al. 2019, 33) and highlighted 'what the value-added is of having youth at the table as equal partners, including *young and emerging evaluators*', particularly in the era of the SDGs (Montrosse-Moorhead et al. 2019, 35, emphasis ours).

Since the International Year of Evaluation in 2015 and the launch of the Global Evaluation Agenda in 2016 (EvalAgenda2020), YEEs have fought for and won space in the global evaluation community and have increasingly contributed to shaping the global evaluation culture (Montrosse-Moorhead et al. 2019; Bennani and Hoosen 2020)¹. Today, they constitute a solid, vital pool for a demand-and-supply workforce and are eager to support the transformation of evaluation practice and the emergence of a powerful evaluation system (Bennani and Hoosen 2020)². With that in mind, how can YEEs look towards 2030 and use the Decade of Action to become transformative agents of change, aligned with the needs of the implementation, follow-up and review of the 2030 Agenda and its 17 SDGs?

As they embark on their voyage to 2030, we invite YEEs to consider shaping their career pathways through the lens of Inkson's (2004) path metaphor, in which he sees a career as a journey with three aspects: first, as *movement* towards an objectively defined destination (delivering the SDGs) through time (from 2020 to 2030) and space (across organizational

¹ All over the world, YEEs have organized themselves to take ownership of EvalAgenda2020, pushing for greater youth participation in evaluation and advocating for what Montrosse-Moorhead and colleagues refer to as a move from 'evaluation on youth to evaluation *with and by youth*' (2019, 40). In doing so, they have positioned themselves as a new generation of evaluation advocates for a 'global youth-participatory evaluation culture', with the role of YEEs defined as collaborator and co-leader (Bennani and Hoosen 2020, 55). The current Eval4Action Campaign (www.eval4action.org), which the United Nations Population Fund Evaluation Office launched in partnership with EvalYouth and more than a dozen national and regional YEE networks, is a vivid example of YEEs' willingness to promote and shape a strong evaluation culture for 2020 and beyond.

² Montrosse-Moorhead and colleagues (2019, 44) underline that, although in previous decades, a limited number of professionals from North America and Western Europe practiced evaluation, many developing countries are now seeing a significant increase in the number of evaluation professionals, including YEEs, particularly in Sub-Saharan Africa, the Middle East and North Africa, and South America. In these regions, YEEs bring cutting-edge technology, connectivity and communication skills to evaluation practice and are eager to enhance their capacities. For example, half of the applications that the EvalYouth Global Mentoring Program receives come from YEEs from Africa alone.

and geographical settings); second, as the *pursuit* of transformational evaluation through processes of continuous learning and change, involving exploration and experimentation to uncover the unknown and make it known; and third, as an *adventure* into uncharted territory, with the goals of enjoying the overall route and learning from paths travelled, fellow travellers and surprises encountered along the way.

In this chapter, we follow in the wake of Montrosse-Moorhead and colleagues by drawing on key takeaways of the Prague Assembly to guide YEEs as they embark on the journey to 2030³. We begin by presenting the notion of ‘professional identity work’ and some of its implications for YEEs who enter the field of evaluation practice. We then propose a framework describing several evaluator identities that YEEs could explore throughout their transformative professional journeys to 2030 – through formal employment engagement and through involvement with voluntary organizations for professional evaluation (VOPEs). We conclude by introducing the concept of ‘identity workspaces’ to suggest that a function that VOPEs and YEE networks could be called upon to fulfil by 2030 is hosting their youngest members’ identity work.

Professional Identity Work in Evaluation: Instrumental and Exploratory Approaches

Shaping one’s career pathway usually implies engaging in ‘professional identity work’, through which professional possible selves can be formed, crafted, maintained, strengthened, revised and reinvented, both in the present and prospectively in the future (Petriglieri, Petriglieri and Wood 2018; Markus and Nurius 1986). Although the literature on career and work role transitions underlines that professional identity work is an ongoing process, it also highlights that it tends to be undertaken most intensely and consciously during specific career junctures, in situations in which individuals transition into a new professional role or field (Petriglieri and Petriglieri 2010).

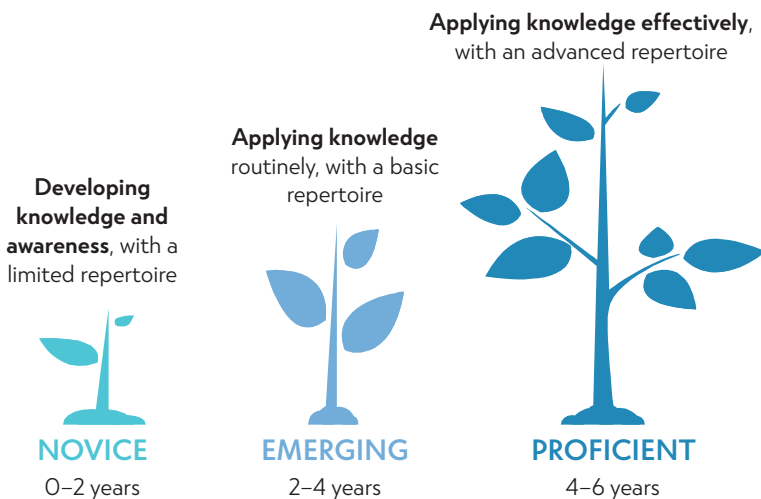
This is precisely the configuration in which most novice evaluators (and thus YEEs) usually find themselves (Stevahn et al. 2005). Per EvalAgenda2020’s definition, YEEs are young evaluators under the age of 35,

³ In particular, we rely on key takeaways from four sessions held under the *Professionalization* strand of the Prague Assembly: special session on YEEs and EvalAgenda2020, workshop on skills assessment and professional development plan, roundtable on career pathways in international development evaluation, session on strengthening the capacities of YEEs in Latin America.

new evaluators with less than five years of experience in evaluation, recent university graduates who are willing to join the evaluation profession or development professionals who have technical knowledge of evaluation and are willing to become evaluation professionals – or some combination of any or all of these (EvalPartners 2016). The material that the EvalYouth Global Mentoring Program uses stresses that, when they join the evaluation community, YEEs enter a robust field that requires acquisition of specific knowledge, skills and dispositions. Their professional growth and development entail a continuous learning process over several years of practice, from novice to emerging to proficient (figure 9.1).

This rather linear approach to professional development and the conduct of identity work can be labelled 'instrumental' (Petriglieri, Petriglieri and Wood 2018, 12). It is centred on delivering performance, understood as the capacity to act credibly and competently in a role. This approach translates into a range of activities aimed at facilitating acquisition and demonstration of knowledge and skills and at pushing YEEs to think and act like competent evaluators in a range of circumstances.

Figure 9.1 Professional Development Process for Novice Evaluators: Instrumental Approach (Delivering Performance)



Source: Bennani (2019). Design based on Stevahn et al. (2005) and material from the 2017 EvalYouth Global Mentoring Program, which was used to facilitate the joint IDEAS-EvalYouth workshop on skills assessment and professional development held at the 2019 Prague Assembly.

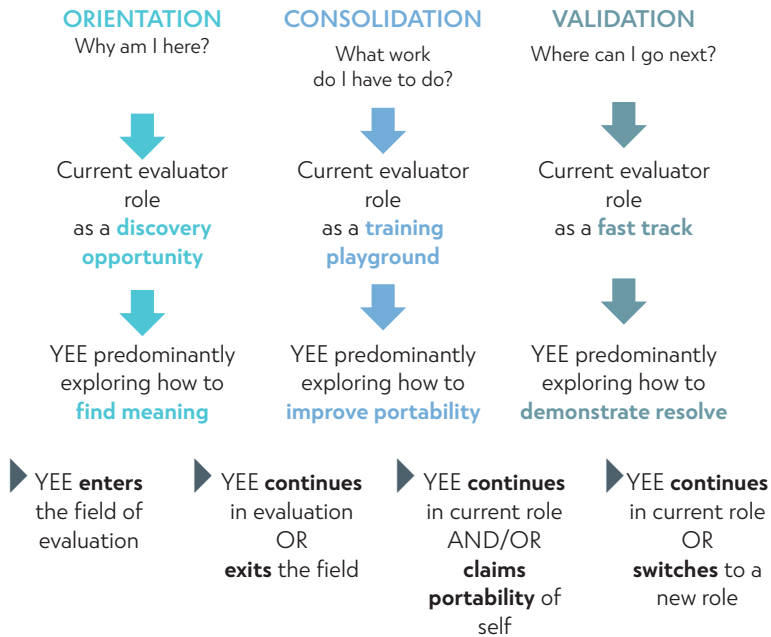
Another way to look at the professional development process of novice evaluators would be to adopt a dynamic lens with an 'exploratory' approach (Petriglieri, Petriglieri and Wood 2018, 13). This approach is centred on finding meaning, understood as the ability to make sense of one's own behaviour and social context. It translates into a range of activities designed to push YEEs to operate reflectively and purposefully in order to examine the meaning and motives of their career choices. It invites them to engage actively in the enterprise of discovering, articulating and shaping their unique professional identities in order to recognize, develop and exercise their capacity to lead in the evaluation world. This approach is based on conceptions of professional development as being experimental and involving identity work and identity play (Ibarra and Petriglieri 2010).

As Ibarra and Petriglieri (2010, 10) put it, 'in organizational life, people work at being certain things but play at becoming others'. Although identity work fundamentally seeks compliance with externally imposed image requirements, identity play is concerned with inventing and reinventing one's professional self. More specifically, identity play is defined as 'people's engagement in provisional but active trial of possible future selves' (Ibarra and Petriglieri 2010, 10). Complementary to the notion of identity work, identity play is a concept that provides a useful starting point to understanding and unveiling the process of discovery necessary for creating one's professional identity. Once identity is in play (open to question and change), a playful posture (as opposed to a rational or efficient one) facilitates exploration and discovery.

In figure 9.2, we present three situational conditions that can foster the dynamics of YEEs' identity construction throughout their professional journeys to 2030. We invite YEEs to approach it as a (perfectible) tool that can help them make sense of the past, make decisions for the present and make plans for the future.

The 'orientation' phase unfolds as YEEs enter the field of evaluation and encounter evaluation practices and fellow evaluators. During this phase, they focus on scanning their environment, understanding the meaning and relevance of their career choice and discovering different types of evaluator roles. They also attempt to figure out ways to use the knowledge and skills they have acquired and apply them to evaluation. When they enter the field of evaluation, YEEs usually have a more or less stable narrative about why they chose to do so. The orientation phase ends when YEEs have a relatively firm answer to the question 'Why am I here?' Depending on the answer, they will choose to keep the narrative going and remain in evaluation or change the plot and exit the field.

Figure 9.2 Professional Development Process for Novice Evaluators: Exploratory Approach (Finding Meaning)



Source: Designed by the authors based on the concepts presented in Ibarra and Petriglieri (2010) and Petriglieri, Petriglieri and Wood (2018).

YEEs who choose to embrace evaluation enter a new 'consolidation' phase, which involves engaging in cycles of social observation, experimentation and self-reflection. YEEs discover more about evaluation practices and evaluator roles. With support and feedback from fellow evaluators, they identify and define the work they need to do to close the gap between their current and desired future selves. The consolidation phase ends when YEEs have a relatively firm answer to the question 'What work do I need to do?' Depending on the answer, they will choose to continue in their current evaluation role and consolidate their skills or prepare for a transition to another role.

YEEs who are interested in switching from their current evaluation role to another enter yet another new phase. During this 'validation' phase, they start demonstrating their resolve to gain recognition of their skills and status as evaluators outside of their current professional configurations. Gaining such recognition is the central concern of this last phase, which

ends when YEEs have a relatively firm answer to the question 'Where can I go next?' Depending on the answer, they will continue in their current evaluation role or transition to a new one.

Ultimately, YEEs' ability to switch from one evaluation role to another will depend on their ability to craft 'portable selves', defined as 'selves endowed with definitions, motives and abilities that can be deployed across roles and organizations over time' (Petriglieri, Petriglieri and Wood 2018, 1). These types of professional identities are fit for itinerant careers that unfold across organizations, sectors and locations, a type of career that is held in high regard today and viewed as a prerequisite for developing the perspective and skills necessary to operate effectively in a globalized world.

In the next section, we present and describe a few evaluator identities that YEEs could explore throughout their transformative professional journeys to 2030.

Departing from Current Professional Configurations: Avenues for YEEs' Journeys to 2030

To assist YEEs in becoming transformative agents of change by 2030, we invite them to consider crafting and shaping their own customized career pathways by combining professional development goals with personal commitments to contribute to the global evaluation community⁴.

The umbrella framework presented in this section (tables 9.1 and 9.2) offers concrete avenues to help our readers conceptualize their career trajectories while also reflecting on ways to build and consolidate transformational evaluation approaches fit for the complexity of the many pressing challenges of our times. This framework attempts to define and differentiate work experiences that YEEs could explore within formal employment engagement or on a voluntary basis through involvement with VOPEs. It offers several prototypes – or role models – of possible selves to supply practical raw material for YEEs' identity work and play

This framework should be considered not as a structured or pre-designed linear career path, but rather as a way for YEEs to *think* of their

⁴ More broadly, this is an invitation to embrace a mandate aligned with the sustainable development ideals of our times. This mandate involves shaping the meaning and exercise of leadership in evaluation as one that encompasses a strong sense of responsible global citizenship.

journeys in order to proactively craft, construct, self-manage and self-direct their own career paths. It aims to provide *inspiration* for how specific roles could be applied throughout the Decade of Action.

Avenues for Work Experiences Within Formal Employment Engagement

At the professional level, there are multiple options for the definition of career paths in evaluation, with different evaluator profiles towards which novice professionals can plan to gravitate. In table 9.1, we deconstruct several professional identities to guide our readers in their exploration. As YEEs embark on their transformative professional journeys to 2030, they may want to reflect on which path will better reflect their skills, ambition and own potential to contribute to the development of transformational evaluation (box 9.1).

Box 9.1 Evaluation Professionals

'One hallmark of professional effectiveness is continuous learning and skills refinement' (Stevahn et al., 2005, 46). All evaluators are learners and should demonstrate specific dispositions to develop and grow. They must be reflective and thoughtful about their own practice in order to strengthen their knowledge and skills. They should be willing to maintain an attitude of personal responsibility towards their development, enhance their motivation and ability to learn from ongoing experiences, and acquire new resources to succeed in the evaluation career.

Avenues for Work Experiences Through Involvement with VOPEs and YEE Networks

At the personal level, there are multiple options for YEEs to commit to the global evaluation community to contribute to the emergence of a powerful evaluation system that can help accelerate and support transformational development.

Over the past decade, the number of national and regional VOPEs has expanded dramatically – from 15 in the 1990s to some 140 nowadays (IOCE

Table 9.1 Roles for Work Experiences Within Formal Employment Engagement

	Evaluation manager	Evaluation practitioner	Evaluation communicator	Evaluation advocate	Evaluation researcher	Evaluation professor
Descriptor: Most prominent differentiator	Commissions and oversees evaluations	Plans, designs and conducts evaluations	Enhances use of evaluations through appropriate evaluative knowledge delivery	Publicly supports need for influential evaluation and evidence-based policies	Produces knowledge on evaluation theory and practice	Teaches evaluation theory and practice
Summary quotation: What matters most in this role	'I secure the production of evaluations that meet the needs of intended users.'	'I conduct evaluations that meet the needs of my clients.'	'I effectively manage, communicate and disseminate evaluation results.'	'I champion stronger evaluation systems, capacities and cultures.'	'I effectively manage, communicate and disseminate evaluation results.'	'I effectively transfer evaluation knowledge and skills to my students.'
Professional background: Typical positions	<ul style="list-style-type: none"> Evaluation officer Senior evaluation officer Head of evaluation service 	<ul style="list-style-type: none"> Evaluation analyst Junior evaluator Senior evaluator Project manager 	<ul style="list-style-type: none"> Editorial expert Communications officer Knowledge manager 	<ul style="list-style-type: none"> Evaluation manager Head of evaluation service parliamentarian Civil society activist 	<ul style="list-style-type: none"> PhD candidate Research fellow Research associate 	<ul style="list-style-type: none"> Graduate student Associate professor Tenured professor Instructor or lecturer
Professional settings: Typical work-places	<ul style="list-style-type: none"> Government Multilateral agency Foundation NGO 	<ul style="list-style-type: none"> Consulting firm University Research Centre (external) Government Multilateral agency Foundation NGO (internal) 	<ul style="list-style-type: none"> Consulting firm Government Multilateral agency Foundation NGO Research centre 	<ul style="list-style-type: none"> Government Multilateral agency Foundation NGO Parliament 	<ul style="list-style-type: none"> University Research centre Think tank NGO 	<ul style="list-style-type: none"> University Training centre
Competencies: Indispensable strengths	<p>Framing evaluation purpose, scope and questions for production of relevant terms of reference</p> <p>Facilitating work of multidisciplinary teams that include highly skilled technical experts</p>	<p>Applying evaluation approaches, methods and tools while maintaining ethical oversight</p> <p>Building strong client relationships while demonstrating commitment to exceptional service</p>	<p>Developing multimedia communication packages comprising relevant, diversified and innovative evaluation content built around key messages and visual assets</p> <p>Making evaluative knowledge accessible to internal and external stakeholders through efficient, user-tailored knowledge management platforms</p>	<p>Shaping narrative and public messaging around key advocacy priorities</p> <p>Nurturing network of influential leaders, powerful decision makers and key stakeholders</p>	<p>Developing profound knowledge of evaluation methods, theories and approaches</p> <p>Conducting qualitative and quantitative research on evaluation theories, methods and approaches</p>	<p>Developing profound knowledge of evaluation methods, theories and approaches</p> <p>Imparting knowledge pedagogically by developing learner-friendly teaching designs</p>

(continued)

Table 9.1 Roles for Work Experiences Within Formal Employment Engagement *(continued)*

	Evaluation manager	Evaluation practitioner	Evaluation communicator	Evaluation advocate	Evaluation researcher	Evaluation professor
Motivations: Key goals	Ensuring production of robust evaluations that meet high standards and requirements Promoting culture of accountability, learning and use of evaluation	Becoming actively involved in 'doing' evaluation, by conducting fieldwork and other reviews Evaluating different units of analysis in different thematic areas	Incorporating culture of communications and knowledge management at each step of evaluation process Bringing the right evidence to the right people at the right time so that it supports decision-making, accountability and learning	Influencing ecosystems by generating increase in demand for and supply of evaluations Generating widespread recognition of evaluation being critical to and a key accelerator for achieving sustainable development	Publishing research results in journals and books Sharing and disseminating generated knowledge at conferences and through webinars	Ensuring that students gradually develop enough knowledge, skills and dispositions to embrace evaluation profession Instilling evaluation vocations in minds of students
Frustrations: Key challenges	Managing potential upfront risks of evaluation findings not confirming existing beliefs and public positions or not being used by stakeholders	Complicated client settings, with limited resources to conduct an evaluation and lack of reliable data to support analysis	Limited interest in evaluations from key stakeholders, resulting in wasted investment and missed opportunity for learning and improving performance	Limited capacity to measure effective results of advocacy efforts as they require long-term strategies	Limited access to real-life evaluations	Limited teaching opportunities, particularly in regions where evaluation training and capacity development programs are not developed
Prospective lens Potential contribution to transformation of evaluation practice during Decade of Action (2020–2030)	Striving to shift commissioning practices from 'business as usual' approach to innovative, transformational approach aligned with foci and needs of Agenda 2030 and its 17 SDGs	Supporting transformation of evaluation practice to be in position to supply evidence-based evaluations aligned with foci and needs of Agenda 2030 and its 17 SDGs	Moving from effectively disseminating evaluation products to strategically communicating evaluative knowledge to different audiences while relying on mechanisms and tools aligned with foci and needs of Agenda 2030 and its 17 SDGs	Changing mindsets and behaviours to support demand and supply of influential, equity-focused, gender-responsible, transformational evaluations aligned with foci and needs of Agenda 2030 and its 17 SDGs	Developing innovative, transformational evaluation theories, approaches and practices aligned with foci and needs of Agenda 2030 and its 17 SDGs	Expanding number of capable evaluators who can fill an important societal role and advancing worldwide spread of evaluation training and practices aligned with foci and needs of Agenda 2030 and its 17 SDGs

Source: Designed by the authors based on their knowledge and understanding of evaluator roles and on IDEAS (2012) and UNFPA Evaluation Office (2019).

Note: NGO = non-governmental organization.

2020; Rugh and Segone 2013). These formally organized societies or associations not only work to increase the supply of high-quality, credible, useful evaluations, but also attempt to address the demand side, by advocating for government policies and systems centred around accountability, learning and public transparency (Kosheleva and Segone 2013). VOPEs offer multiple ways for YEEs all over the world to engage meaningfully in volunteer work in order to share experiences, learn from seasoned peers, plan their professional development, discuss the future of the field, form partnerships to advance the evaluation profession and advocate for demand and supply of transformational approaches to evaluation.

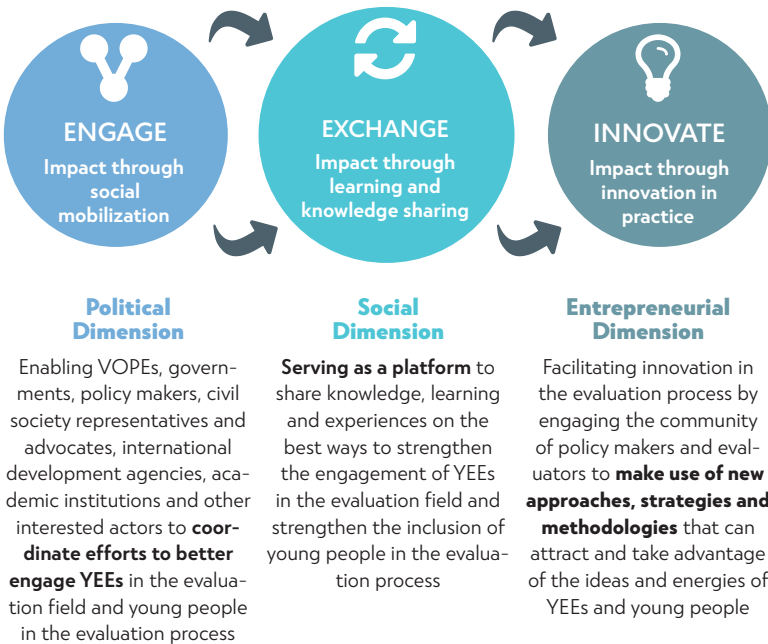
In addition to VOPEs, we have seen the launch of multiple national and regional YEE networks in the Global North and South⁵. These networks have been cooperating in unprecedented ways under the global umbrella network EvalYouth, an initiative that EvalPartners launched in 2015 to promote engagement, exchange and innovation among YEEs, youth stakeholders and other key actors. All over the world, YEE networks have taken full ownership of EvalAgenda2020, operating through action plans aligned with the three pillars described in figure 9.3 (EvalPartners 2016).

VOPEs and YEE networks are two types of fora that offer the possibility of engaging across boundaries and encountering diversity by working with evaluation peers from different backgrounds, sectors and regions.

The United Nations (2016) public outreach website dedicated to the SDGs lists and defines five potential roles young people (in general) can play in the overall effort to achieve the SDGs (United Nations 2016): leaders, change-makers, innovators, critical thinkers and communicators. While reflecting on ways to commit as volunteers through involvement with VOPEs and YEE networks, YEEs could find interesting ways to apply these five youth roles in the evaluation field. They could consider them to define the mindset and position with which they intend to engage and contribute to the global evaluation community as transformational evaluators. Table 9.2 provides a few details on these roles, as well as some examples that illuminate how YEEs could perform them throughout their transformative professional journeys to 2030 (box 9.2).

⁵ Regional networks include EvalYouth Latin America and the Caribbean, EvalYouth Europe and Central Asia, EvalYouth Asia, EvalYouth Middle East and North Africa, Francophone Network of YEEs, European Evaluation Society YEEs and African Evaluation Association YEEs. In many countries, they are often supplemented by national YEE networks affiliated with national VOPEs. At the global level, IDEAS has also recently integrated a YEEs Thematic Interest Group within the scope of its work.

Figure 9.3 EvalAgenda 2020 Pillars for Cultivation of Novice Evaluators' Individual Capacities



Source: Designed by authors based on EvalPartners (2016, 63–69).

Box 9.2 Evaluation Ambassadors

'When you label yourself (or others label you) as an evaluator, you represent us all' (Podems 2019, 216). By undertaking work experiences with VOPEs and YEE networks, YEEs can fulfil the overarching role of 'evaluation ambassador', which entails advocating for the institutionalization of evaluation; promoting the worldwide recognition of evaluation as a profession; educating others on evaluation standards and ethics and raising awareness of evaluation societies, groups and other organizations that are relevant to the contexts in which evaluators work.

Table 9.2 Roles for Work Experiences Through Involvement with Voluntary Organizations for Professional Evaluation and Young and Emerging Evaluator Networks

<p>Descriptor: Most prominent differentiator</p>	<p>Evaluation Leader</p>	<p>Evaluation Change Maker</p>	<p>Evaluation Innovator</p>	<p>Evaluation Critical Thinker</p>	<p>Evaluation Communicator</p>
<p>Summary quotation: What matters most in this role</p>	<p>Drives change in their community, country, region</p>	<p>Believes in power to act and mobilize others</p>	<p>Brings fresh perspectives and offers new ideas and alternative solutions</p>	<p>Identifies and challenges existing power structures and barriers to change</p>	<p>Communicates development agenda to peers and communities</p>
<p>Context: Configuration in which role will be performed</p>	<p>'I strive for the development of strong national and regional evaluation capacities'. VOPEs and YEE networks are often managed by boards and management groups that rely on commitment and work of highly motivated volunteers YEEs are still under-represented in leadership, governance and decision-making bodies of many VOPEs.</p>	<p>'I actively work for the establishment of a strong enabling environment for evaluation in my region or country'. Despite its growing acceptance, evaluation has not been embraced as widely as it should be. There is inadequate appreciation of what it is, how it differs from policy research and performance auditing and measurement, and how it can help improve policymaking and programme implementation efforts on a practical level.</p>	<p>'I contribute to thought leadership for development of new approaches to evaluation'. VOPEs often host thematic working groups that seek to explore reframing of evaluative practice around the 2030 Agenda and its 17 SDGs. Some networks such as the SDG Transformations Forum can also offer opportunities to contribute to development of alternative solutions.</p>	<p>'I raise awareness of the need for equity-focused, gender-responsive evaluation systems'. To facilitate transformative change, policymakers need equity-focused, gender-responsive evaluative evidence. There is a need for appropriate participation from all stakeholders, including youth, who are frequently absent from evaluation processes that affect them.</p>	<p>'I join advocacy coalitions to promote influential evaluation as an essential management and social change instrument'. Youth activists and YEEs are engaging all over the world to promote the need to achieve the SDGs and call for a stronger consideration of them in national plans and policies, including by conducting voluntary national reviews. Gaining access to broader connectivity and use of social media is bolstering their efforts</p>
<p>Contributions: Typical actions to undertake</p>	<p>Improving YEEs' representation by running for board elections and taking the lead on flagship initiatives to bring the specific perspectives of youth into the work of VOPEs</p>	<p>Advocating for achievement of a positive evaluation culture that embeds broad-based evaluative thinking and drives policies geared to the public interest</p>	<p>Engaging in projects that target transformation of evaluation practice by focusing on development of new approaches such as Blue Marble Evaluation*</p>	<p>Providing policymakers with guidance on use of transformative approaches to evaluation that can help identify structural causes of inequalities through deeper analysis of power relationships, social norms and cultural beliefs</p>	<p>Launching and leading advocacy initiatives such as the Eval4Action Campaign* to influence policymakers and other key stakeholders so public policies are based on evidence and incorporate considerations of equity and effectiveness</p>

(continued)

Table 9.2 Roles for Work Experiences Through Involvement with Voluntary Organizations for Professional Evaluation and Young and Emerging Evaluator Networks *(continued)*

	Evaluation Leader	Evaluation Change Maker	Evaluation Innovator	Evaluation Critical Thinker	Evaluation Communicator
Contributions: Typical actions to undertake <i>(continued)</i>	Strengthening the institutional capacities of VOPEs and YEE networks and supporting their vision and dynamism over time to advance the interests of the evaluation profession at the national and regional levels Building a service line that can enlarge the pool of skilled evaluators to secure production of high-quality transformational evaluations	Building relationships with members of the political and policy sphere to help strengthen the capacity to demand and use evidence generated by evaluations to inform policy decision-making and implementation Building relationships with civil society organizations to promote evaluation as an instrument of transparent, accountable, collaborative, inclusive governance	Publishing innovative ideas and approaches in evaluation journals and other official publications Sharing and discussing innovative ideas and approaches during evaluation workshops, conferences and webinars	Disseminating knowledge on the use of mixed-methods, big data and new information technology for assessing complexity, causality, sustainability and resilience Advocating for a global youth participatory evaluation culture that positions youth and YEEs as partners and co-leaders and fully integrates their energy, perspectives, values and skills	Positioning YEEs as partners of other young people by acting jointly with youth activist networks to advocate for young people to be more systematically involved in implementing and reviewing progress towards the SDGs Connecting to voluntary national review processes and other local, national or regional public mobilization campaigns on SDG implementation, gender equality, climate change etc., to advocate for use of evaluation in policymaking
Prospective lens: Potential contribution to transformation of evaluation practice during Decade of Action (2020–2030)	Shifting service line and advocacy work of VOPEs and YEE networks from a business-as-usual approach to an innovative, transformational approach aligned with foci and needs of Agenda 2030 and its 17 SDGs	Contributing to emergence of powerful global community of VOPEs and YEE networks able to influence policymakers, public opinion and other key stakeholders in a way that is aligned with foci and needs of Agenda 2030 and its 17 SDGs	Advancing transformational evaluation knowledge, methods and tools to lead to breakthroughs aligned with foci and needs of Agenda 2030 and its 17 SDGs	Ensuring systematic use of transformational equity-focused, gender-responsive evaluations aligned with foci and needs of Agenda 2030 and its 17 SDGs	Disseminating knowledge about and promoting use of strong, influential transformational evaluation approaches aligned with foci and needs of Agenda 2030 and its 17 SDGs

source: Designed by the authors based on their own knowledge and understanding of evaluator roles and on Action Aid (2020); Bamberger, Segone and Tateossian (2016); Ben Salem, Mouafo Nopi and Dieffi Tchifou (2020); De Sole (2018); Karkara et al. (2014); Olavarria and Sánchez-Romero (2019); UN DESA (2015, 2018); Velez and Alkowni (2018).

• Blue Marble Evaluation is a global initiative launched by Utilization-Focused Evaluation designed to train the next generation of evaluators in global systems evaluation (Patton 2019).

• Eval4Action, also known as the Decade of Evaluation for Action campaign, is a call for stronger evaluation capacities to accelerate progress towards the SDGs that the United Nations Population Fund Evaluation Office, EvalYouth, and the Global Parliamentarians Forum for Evaluation launched in April 2020 and some 115 international and national institutional partners have joined (UNFPA Evaluation Office 2020).

Looking Forward: Towards the Emergence of Evaluation 'Identity Workspaces'?

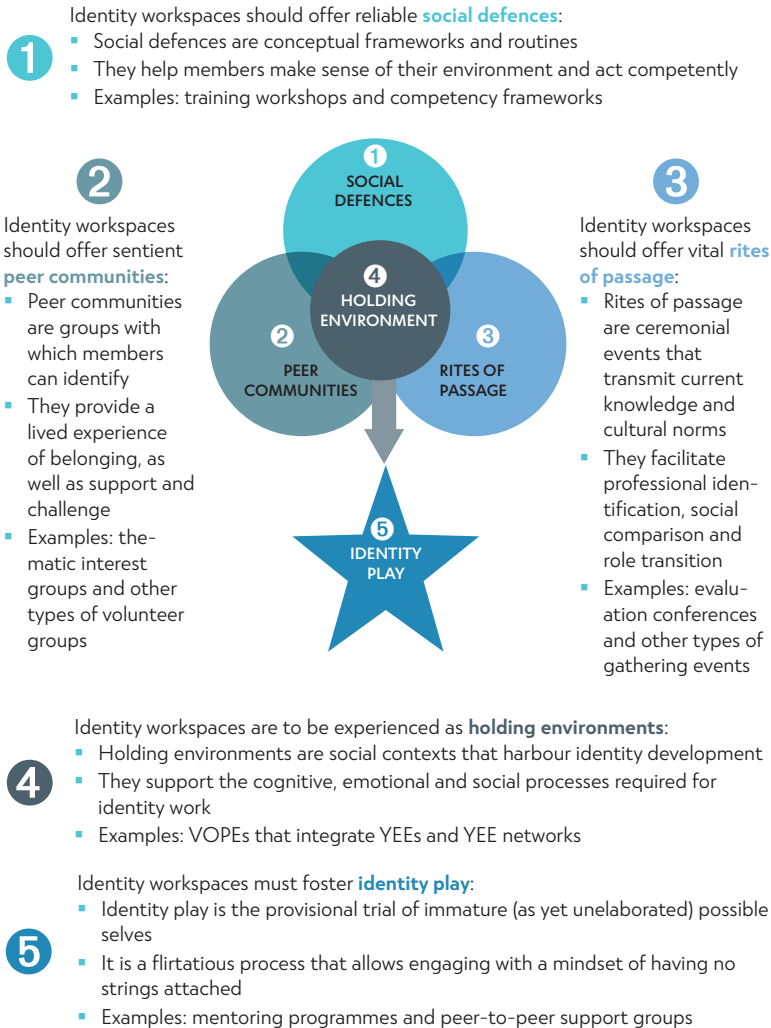
In this last section, we introduce the concept of 'identity workspaces', defined as 'institutions that provide a holding environment for individuals' identity work' (Petriglieri and Petriglieri 2010, 44). We posit that a function that VOPEs and YEE networks could be called upon to fulfil by 2030 is hosting their youngest members' identity work. By serving as identity workspaces, VOPEs and YEE networks would go beyond influencing what the new generation of evaluators knows and does by supporting YEEs in discovering, understanding and shaping who they are and can become in the era of the SDGs – or, better yet, who they are *becoming*. As such, YEEs could consider them as physical, social and psychological spaces that offer the possibility of contextualizing and personalizing their transformative journeys to 2030 in order to constantly revise, consolidate and redraft their professional narratives as transformational evaluators in the making.

As shown in figure 9.4, institutions are trusted as identity workspaces when they offer a combination of five features. Each of these features sustains the psychological and social adjustments underpinning professional identity work.

The literature on career and work role transition emphasizes the importance of having sentient peer communities and stresses that exploring possible identities is not an activity that can occur in isolation (Petriglieri and Petriglieri 2010; Kreiner, Hollensbe and Sheep 2006). To be entrusted with the function of identity workspaces, VOPEs and YEE networks should offer opportunities to sustain connections and foster identification. The role that referent others⁶ fill in identity work is essential and often decisive (Schwartz and Ames 1977). The guidance of seasoned professionals and peers can shape YEEs' expectations about the range of open possibilities and encourage them to delve into different alternatives. Counsellors and mentors can also play a central role, facilitating the effective practice of YEEs as they craft, construct and assemble their careers. In this regard, two initiatives that YEE networks have launched are worth highlighting as relevant examples that could inspire future endeavours for the facilitation of YEEs' professional identity work by 2030. In table 9.3, we briefly present these two initiatives: the EvalYouth Global Mentoring Programme, launched

⁶ 'Referent others' are individuals or groups who serve as role models and have a particularly strong influence on the process of self-definition.

Figure 9.4 Identity Workspaces for the Professional Development of Young and Emerging Evaluators



Source: Designed by authors based on the concepts presented in Petriglieri and Petriglieri (2010) and Ibarra and Petriglieri (2010).

in 2017, and the Peer-to-Peer Career Advisory Sessions for YEEs, launched in 2019. We believe that they each present interesting features that could be embedded in the work of VOPEs and YEE networks willing to become well-suited enough to serve as identity workspaces.

Table 9.3 Mentoring and Peer-to-Peer Support Initiatives for the Professional Development of Young and Emerging Evaluators

Element	EvalYouth Global Mentoring Programme	Peer-to-Peer Career Advisory Sessions for YEEs
Hosting	Hosted exclusively by EvalYouth Global (Task Force 2)	Hosted jointly by several YEE networks (European Evaluation Society YEEs, EvalYouth, IDEAS, JEEunes)
Focus	Elder focused: matches YEEs with seasoned evaluators	Peer focused: connects YEEs with YEEs
Format	Mentees follow online learning modules and hold regular meetings with mentors to discuss level of attainment	Online focus group meetings held on virtual conference platform with rounds of questions and answers between YEE participants and two YEE facilitators
Purpose	Provide a formal curriculum to acquire knowledge and skills, identify and overcome challenges and barriers, and formulate actionable professional development plans	Provide informal safe space where YEEs can share sensitive work challenges and exchange advice
Philosophy	Mentor/mentee pairs are matched based on mutual professional interests and regions of work	Content of sessions is bottom up and demand driven, and facilitators adopt nonjudgemental mindset

Source: Authors, based on material used by EvalYouth Global Mentoring Program and Peer-to-Peer Career Advisory Sessions for YEEs concept paper.

Note: According to the needs assessment that EvalYouth conducted in 2016, mentoring and understanding of career path were indicated as the most important unmet needs of YEEs. YEE networks have been actively working in response to this need.

Both initiatives rely on the assumption that an evaluator's professional identity is a project to be worked on and that identity play requires a relatively safe space to try out new and untested behaviours. The EvalYouth Global Mentoring Program is primarily centred on the instrumental approach to professional development, with a strong focus on acquisition of knowledge and skills through online learning modules and a defined framework for competency development, but the programme also makes room for play; the bond that can develop between the YEE and the guiding mentor forms a transitional space within which the fantasy of a future possible self can start becoming a reality. Conversely, the Peer-to-Peer Career

Advisory Sessions are primarily centred on the exploratory approach to professional development. With a strong focus on the discovery of alternatives, these sessions can shape YEEs' expectations about the range of possibilities open to them and motivate them to pursue roles in the field of evaluation and eventually become well-established, seasoned evaluators.

Ultimately, VOPEs and YEE networks will be well suited to serving as identity workspaces for YEEs in the era of the SDGs when they include a significant experimental component, combine the acquisition of knowledge and skills with opportunities for personal reflection and experimentation, inspire YEEs to pursue long-term development, involve learning about the activities and identities associated with evaluation practice, expose YEEs to the interplay between individual functioning and the group dynamics of the global evaluation community and foster the strength of a shared culture. More importantly, to be truly considered identity workspaces by YEEs throughout their transformative professional journeys to 2030, VOPEs and YEE networks should provide safe training grounds in which YEEs can experiment freely, with features similar to those of formal work environments but more forgiving.

Conclusion

The Decade of Action offers YEEs an opportunity to become major contributors to development of transformational evaluation approaches fit for the complexity of the many pressing challenges of our times. In this chapter, we have invited them to depart from their current professional configurations to embark on transformative professional journeys that can take multiple forms, with combinations of complementary work experiences acquired in more (professional engagement) and less (volunteer engagement) formal settings. To be truly meaningful and valuable for YEEs, such journeys to 2030 cannot be undertaken alone. Institutional and individual partners, with which YEEs should learn to build and nurture strong relationships, constantly facilitate careers in evaluation. We also invite our readers to approach their journeys as an opportunity to revisit the motif of Campbell's hero myth. As Christopher Vogler (2007, 7) puts it, 'every storyteller bends the mythic pattern to his or her own purpose or the needs of a particular culture, that's why the hero has a thousand faces'. As they embark on the road to 2030, YEEs should absorb the ideas and avenues presented in this chapter; recreate them with fresh insights and personal sense of self and draft their own, unique professional narratives as transformational evaluators in the making.

References

- ActionAid. 2020. *Believe in better. A working paper on young people's inclusion in national follow-up, review and accountability processes of the 2030 Agenda for sustainable development*. Johannesburg: ActionAid.
- Allison, Scott, and George Goethals. 2017. The hero's transformation. In: *Handbook of heroism and heroic leadership*, (eds.) Scott Allison, George Goethals and Roderick Kramer, 379–400. New York: Routledge.
- Bamberger, Michael, Marco Segone and Florencia Tateossian. 2016. *Evaluating the Sustainable Development Goals with a 'no one left behind' lens through equity-focused and gender-responsive evaluations*. UN Women, EvalGender and IOCE in partnership with EvalPartners, UNEG, UNICEF, UNFPA and Swiss Agency for Development and Cooperation, New York.
- Ben Salem, Sana, Arnoux Mouafo Nopi and Miltiade Dieffi Tchifou. 2020. *The success stories of the Francophone Network of Emerging Evaluators*. EvalYouth.
- Bennani, Kenza. 2019. Designing a 'Made in Africa Evaluation' competency framework for young and emerging evaluators. Poster presented at the 9th AfrEA Conference. Abidjan.
- Bennani, Kenza, and Fazeela Hoosen. 2020. A new generation of advocates: Young and emerging evaluators and the promotion of a global youth-participatory evaluation culture. In *Promoting an evaluation culture in 2020 and beyond. Evaluation matters*. Q1: 54–63. Independent Development Evaluation, African Development Bank Group, Abidjan.
- Campbell, Joseph. 1949. *The hero with a thousand faces*, 3rd ed. New York: New World Library.
- De Sole, Natalie. 2018. *EvalYouth International Mentoring Program. Pilot phase external evaluation*. Rooted-Growth, EvalYouth and EvalPartners. <https://evalpartners.org/evalyouth/mentoring-program>
- EvalPartners. 2016. *Global Evaluation Agenda 2016–2020*. <https://www.evalpartners.org/global-evaluation-agenda>
- EvalYouth Global. 2017. *Material of the Global Mentoring Program*. Used to facilitate the joint IDEAS-EvalYouth workshop on skills assessment and professional development held at the Prague Assembly in 2019.
- Ibarra, Herminia, and Jennifer Petriglieri. 2010. Identity work and play. *Journal of Organizational Change Management* 23 (1): 10–25.
- IDEAS (International Development Evaluation Association). 2019. *Prague Declaration on Evaluation for Transformational Change*. Adopted at the 2019 IDEAS Global Assembly and Third International Conference on Evaluating Environment and Development. <https://tinyurl.com/bcm294k7>
- IDEAS. 2012. *Competencies for development evaluation evaluators, managers, and commissioners*. IDEAS Working Group on Evaluation Competencies.
- Inkson, Kerr. 2004. Images of career: Nine key metaphors. *Journal of Vocational Behavior* 65: 96–111.

- IOCE (International Organization for Cooperation in Evaluation). 2020. VOPEs Database. <https://ioce.net/vopes>
- Karkara, Neha, Marco Segone, Tessie Catsambas et al. 2014. *Advocating for evaluation. A toolkit to develop advocacy strategies to strengthen an enabling environment for evaluation*. UN Women, EvalPartners and IOCE in partnership with Ministry of Foreign Affairs of Finland, OECD, UNEG, UNICEF, USAID, New York.
- Kosheleva, Natalia, and Marco Segone. 2013. EvalPartners: Facilitating the development of a new model of voluntary organization for professional evaluation to support the development of national evaluation capacities. *American Journal of Evaluation* 34 (4): 568–572.
- Kreiner, Glenn, Elaine Hollensbe and Mathew Sheep. 2006. Where Is the 'me' among the 'we'? identity work and the search for optimal balance. *Academy of Management Journal* 49 (5): 1031–1057.
- Markus, Hazel, and Paula Nurius. 1986. Possible selves. *American Psychologist* 41 (9): 954–969.
- Montrosse-Moorhead, Bianca, Khalil Bitar, Josette Arévalo et al. 2019. Evolution in the making: Evaluation 'done well' in the era of the SDGs with a youth participatory approach. In *Evaluation for transformational change. Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvia Salinas, 33–50. Exeter, UK: International Development Evaluation Association.
- Olavarría, Claudia, and Gerardo Sánchez-Romero. 2019. *What works for VOPEs and YEEs? Quick guide to welcome young and emerging evaluators to organizations*. EvalYouth Latin America and Caribbean.
- Patton, Michael Quinn. 2019. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Petriglieri, Gianpiero, Jennifer Petriglieri and Jack Wood. 2018. Fast tracks and inner journeys: Crafting portable selves for contemporary careers. *Administrative Science Quarterly* 63(3): 479–525.
- Petriglieri, Gianpiero, and Jennifer Petriglieri. 2010. Identity workspaces: The case of business schools. *Academy of Management Learning and Education* 9 (1): 44–60.
- Podems, Donna. 2019. *Being an evaluator. Your practical guide to evaluation*. New York: Guilford Press.
- Rugh, Jim, and Marco Segone (eds.). 2013. *Voluntary organizations for professional evaluation (VOPEs): Learning from Africa, Americas, Asia, Australasia, Europe and Middle East*. New York: EvalPartners, UNICEF, IOCE in partnership with Cooperación Española, Ministry of Foreign Affairs of Finland, UNEG and UN Women.
- SDG Transformations Forum. 2017. *Transforming evaluation*. <https://transformationsforum.net/t-evaluation/>
- Stevahn, Laurie, Jean King, Gail Ghore et al. 2005. Establishing essential competencies for program evaluators. *American Journal of Evaluation* 26 (1): 43–59.

- Schwartz, Shalom, and Ruth Ames. 1977. Positive and negative referent others as sources of influence: A case of helping. *Sociometry* 40 (1): 12–21.
- United Nations. 2016. *Youth and the SDGs*. <https://www.un.org/sustainabledevelopment/youth/>
- UN DESA (United Nations Department of Economic and Social Affairs). 2018. *Youth and the 2030 Agenda for Sustainable Development*. World Youth Report 2018. New York: UN.
- UN DESA (United Nations Department of Economic and Social Affairs). 2015. *Youth civic engagement*. World Youth Report 2015. New York: UN.
- UNFPA (United Nations Population Fund) Evaluation Office. 2020. EvalYouth Global Network. Global Parliamentarians Forum for Evaluation. Decade of evaluation for action to deliver the SDGs by 2030. <https://www.eval4action.org/>
- UNFPA (United Nations Population Fund) Evaluation Office. 2019. *Strategy to enhance evaluation use through communications and knowledge management (2018–2021)*. New York: UN.
- United Nations General Assembly. 2015. *Transforming our world. The 2030 Agenda for sustainable development*. A/RES/70/1.
- Van den Berg, Rob D., Cristina Magro and Silvia Salinas Mulder (eds.) 2019. *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*. Exeter, UK: IDEAS.
- Vélez, Alberto, and Amani Alkowni. 2018. *Highlighted experiences of the project 'Success Stories from EvalYouth Chapters' – EvalYouth Latin America and the Caribbean*. EvalYouth Global.
- Vogler, Christopher. 2007. *The writer's journey. Mythic structure for writers*, 3rd ed. Studio City, CA: Michael Wiese Productions.



CHAPTER 10

Towards an International Evaluation Academy as an Agent of Transformational Change

LINDA G. MORRA IMAS

Abstract. For decades, the value of evaluation professionalization has been debated. A prolific evaluation literature is now available. This chapter puts forward a transformational concept designed in part to promote evaluation professionalization: an international evaluation academy (IEAc). The 2019 International Development Evaluation Association Global Assembly, held in Prague, culminated in approval of a declaration that supported exploration of an IEAc initiative to act as a platform for innovation, creativity and collaboration in pursuit of evaluation professionalism and influence. This chapter summarizes the pros and cons of professionalization; examines responses to an international survey that confirm broad-based support for the IEAc concept, including a focus on professionalization and outlines what the IEAc is about and some ways it will address evaluation professionalization.

Introduction

For several years, the debate on the status of the profession of evaluator has been a driving force in the community of evaluators and has fed controversies between partisans of the adoption of mechanisms and means of professionalization and defenders of the status quo (Jacob and Boisvert 2010).

This familiar-sounding quotation comes from Jacob and Boisvert's seminal synthesis on the professionalization of evaluation published in *Evaluation* in 2010. The article was published after the Canadian Evaluation Association launched its ground-breaking Credentialed Evaluator designation in 2009. More than a decade later, the arguments for and against evaluation professionalization outlined in the article remain current (table 10.1), yet articles and blogs debating the value of evaluation professionalization continue to abound (e.g. Gauthier 2019; Heider 2015; 2018; Morra Imas 2010; 2017; Picciotto 2011; Quesnel 2010; Schwandt 2017; UNEG 2016). In parallel, a broader case for evaluation transformation is being made in response to the Sustainable Development Goals (SDGs) and the COVID-19 pandemic (e.g. Ofir 2020; Patton 2020; Van den Berg, Magro and Mulder 2019). These global challenges raise broader questions about professionalization and call for holistic evaluation approaches that look for interconnections between the global and the local.

This more comprehensive agenda may explain why adoption of the Canadian Evaluation Association's credentialing process by other evaluation organizations has been limited. Similarly, the customized processes of professional development, involving self-assessment and self-reflection regarding competencies, backed up by voluntary peer review processes sponsored by the United Kingdom Evaluation Society and the European Evaluation Society, have not had much traction¹. Evidently, evaluation organizations, even mid-sized ones such as the International Development Evaluation Association (IDEAS), with 400 members, lack the resources to operate a credentialing process.

Although progress on professionalization has slowed since 2010, the visibility of evaluation increased when the United Nations declared 2015 the Year of Evaluation and approved regular tracking of the universally endorsed SDGs. In the same year, a global evaluation agenda (EvalAgenda) was endorsed. EvalAgenda visualized a world in which evaluation would be

¹ The European Evaluation Society suspended its version of the Voluntary Evaluator Peer Review process in 2020.

Table 10.1 Synthesis of Arguments on Evaluation Professionalization

Pros	Cons	Challenges
Strengthening the field and establishing some boundaries for the profession	Homogenizing evaluation and restricting diversity	Difficulty defining the specificities of the field of evaluation
Increasing evaluation training offerings	Reducing training offerings	Difficulty identifying and verifying the expertise and contribution of the evaluator
Enhancing and improving the status and prestige of evaluation	Restricting or blocking access to the profession	Need to create new structures or organizations
Facilitating selection of evaluators and improving quality of conducted evaluations	Turning evaluation in on itself	
Protecting the public		
Avoiding problematic or unethical behaviour		

Source: Jacob and Boisvert (2010).

an integral part of *all* government, civil society and private sector development efforts, and although only 20 evaluation associations were in existence in 2000, the number has since exploded. EvalPartners² has identified 145 active associations or networks, of which 103 are at the national level.

The International Organization for Cooperation in Evaluation (IOCE) was created in 2003 with a mission 'to increase public awareness and globally validate evaluation and support voluntary organizations of professional evaluation (VOPEs) in contributing to good governance, effective decision-making and strengthening the role of civil society'. Its main activity has been to encourage and support organizational capacity building for evaluation organizations. Although the IOCE has been effective in this role, with about half of the countries in the world yet to be served by a national evaluation society, it seemingly still has a large job ahead.

² The IOCE and the United Nations formed EvalPartners, whose members are civil society organizations and VOPEs.

In 2015, IOCE launched a catalytic effort to advance the professionalization of evaluation. The initiative aimed to increase access to information on evaluation professionalization, provide a platform for an inclusive discussion on professionalization and facilitate cooperation on professionalization initiatives. Documents were collected and put online³, but discussion participation and general interest was low. In 2017, the IOCE attempted to revive the programme and formed the IOCE Professionalization Task Force, which met with no more success and concluded that it was too early to aim for a unified approach to professionalization. In line with the EvalA-genda, the focus shifted to supporting individual VOPEs and recognizing their diversity.

In another development, Julnes and Newcomer (2018) proposed sponsorship of a national evaluation institute to the American Evaluation Association. They envisioned it as a vehicle for supplying expert guidance on evaluation to governments, private and civil society organizations and other stakeholders, but the idea has had little take-up. In 2019, at its annual meeting, the association put together a group to discuss global core evaluation standards. Participants agreed on the concept of 60 per cent of evaluation standards being core standards, with the rest leaving room for cultural and organizational differences. Again, follow-up action has been limited.

In 2020, the Asian Pacific Evaluation Association sponsored an inter-regional initiative for the professionalization of evaluation. As part of this project, a survey was conducted of the requisite conditions for professionalization in six countries – India, Sri Lanka, Kenya, Philippines, Kyrgyzstan and Western Balkans. This involved, for each country, systematic examination of institutional mechanisms and systems for evaluation in the public sphere, the existence and role of professional associations of evaluators and facilities for education and training in the field. That the study used this framework in all six countries makes it particularly useful.

The survey findings were encouraging.

- All surveyed countries had evaluation associations, although they were active to varying degrees.
- Most had strong monitoring systems.
- An environment conducive to evaluations was gradually emerging in the surveyed countries, and the demand for evidence was growing.

³ <https://www.ioce.net/professionalization>.

Nevertheless, the overall conclusion was that, in most countries, professionalization of evaluative function and practice, in terms of development of competency frameworks, educational programmes focused on evaluation and implemented national policies on evaluation, was not in evidence. For example, there are no full-fledged academic monitoring and evaluation (M&E) courses in any Indian university leading to a masters or doctoral degree. Sri Lanka has taken steps towards professionalization in having a strong monitoring system, offering a postgraduate diploma in M&E conducted by faculty of graduate studies at the University of Sri Jayewardenepura and an approved national evaluation policy and strong evaluation society, but evaluation has not yet been adopted in the public sector. The Kyrgyz Republic is struggling to develop an evaluation culture, and capacities are low. The Philippines has a national evaluation policy but a weak M&E system with little demand and has no universal guidelines or competencies or specific M&E training. In the Balkan countries, use of evaluation is minimal, and M&E systems are weak, as are evaluation capacities. Kenya is a relatively bright spot with its active, new Evaluation Society of Kenya, and various universities and institutions provide M&E education and training, although its national M&E policy has been in draft form for longer than 10 years. In all, the study reinforced that there is a vast unmet need for evaluation professionalization.

Although the survey covered only six countries, recent reviews of M&E in Africa and Latin America (Gounou and Perez Yarahuan 2019) and in the small island developing states in the Caribbean and the Pacific (Baptiste and Iese 2019) echo the above findings: There is progress. For example, most Latin America governments formally recognize M&E activities. Since the 2010s, countries cooperating in the Twende Mbele programme – Benin, Uganda, South Africa – have made a significant effort to mainstream evaluations in the work of government. Each, for example, has a national evaluation policy. Nevertheless, in most African countries, monitoring is strong, but evaluation systems and processes are often missing or misunderstood for their role in the SDGs. In Latin America, in the past 20 years, capacities have been built, information on public programmes has been gathered, programme logic models have been developed and evaluations have been delivered, but progress has not been homogeneous in terms of consistency and quality across countries, sectors and time. The use of evidence to increase the effectiveness of programmes and policies is weak, and the authors note that progress is politically fragile, citing the case in Mexico of the conditional cash transfer programme, PROSPERA, a heavily evaluated programme that was cancelled despite numerous evaluations showing positive results.

Baptiste and Iese, discussing evaluation in the Caribbean and in Pacific small island developing states, indicate that the first challenges that evaluators working in the two regions face is the scarcity of M&E systems at the project level and nationally. They note the challenge of the limited in-country technical capacity of regional personnel to conduct evaluations.

In all, these studies confirm the conclusion of the Asian Pacific Evaluation Association study: there has been considerable progress, and bright spots along the way, but we are far from evaluation professionalization.

The Prague Declaration

An impetus for the concept of an international evaluation academy (IEAc) was the gathering of evaluators, commissioners, parliamentarians and other evaluation users at the IDEAS Global Assembly, held jointly with the Third International Conference on Evaluating for Environment and Development in the Czech Republic from 30 September to 4 October 2019. At a workshop preceding the conference (Morra Imas 2019) and in conference presentations, the concept of an IEAc was explored. The conference event culminated in approval of the Prague Declaration, which among other things, supported development of an IEAc⁴, as discussed in the last chapter of this volume. As a follow-up to this historic declaration, a note was put forward that broadly outlined preliminary design ideas for an IEAc, taking into account the constructive deliberations that took place during a Global Assembly workshop and a parallel session informed by a concept paper (Morra Imas 2019).

The Prague Declaration expressed strong support for the social transformations needed to help implement the SDGs through participatory evaluation approaches that respect human rights, promote gender equality and live up to the 'leave no one behind' mission of the SDGs. It highlighted the need for evaluators to address the existential threats of climate change and other 'problems without passport' (Annan 2009). Global Assembly participants also concluded that, to live up to its potential in a world in which no individual, no community and no country exists in isolation, evaluation must be transformed through partnerships grounded in mutual trust, shared visions, ethical codes and mutually agreed professional standards. In this context, the declaration endorsed the proposed development of an IEAc committed to advancing professionalization; promoting interactions

⁴ See No. 6 in <https://tinyurl.com/bcm294k7>.

between science, research and evaluation; enriching the evaluation enterprise and intensifying efforts to build evaluation capacities at all levels.

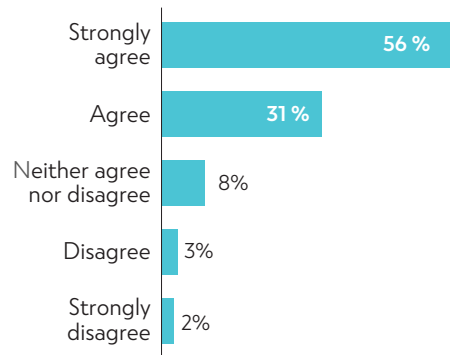
Beyond the Prague Declaration

After the Prague Declaration was issued, questions remained as to the extent of practitioners' demand for an IEAc. This question was highly pertinent given the IOCE's failure to encourage debate and initiate action in support of professionalization. The broader concept of an IEAc along the lines sketched above was tested using a five-question survey that IDEAS sponsored in February 2020. It was sent to IDEAS members and other Global Assembly on Evaluation for Transformative Change and Third International Conference on Evaluating for Environment and Development attendees, as well as International Program for Development Evaluation Training listserv, Facebook and LinkedIn members and IDEAS LinkedIn members, going to an estimated 3,500 individuals. Responses were received from 458 individuals for an approximately 13 per cent response rate⁵.

Overall, the survey disclosed strong support for the concept, with approximately 87 per cent of 454 respondents agreeing (31 per cent) or strongly agreeing (56 per cent) that an IEAc should be set up to complement the work of associations (figure 10.1) and only approximately 6 per cent disagreeing or strongly disagreeing. Extensive comments – about half of the respondents wrote comments – and offers of assistance helped in further planning and designing an IEAc.

Another of the survey questions asked about what functions an IEAc should have if it were set up: support for evaluation professionalism initiatives;

Figure 10.1 Support for an International Evaluation Academy



Note: $n = 452$.

⁵ The full report on the survey can be found on the IDEAS website: www.ideas-global.org (10March2020 IEAc IDEAS Survey Final Summary Report, <https://tinyurl.com/yjnm39nu>).

harmonization of evaluation principles, guidelines and norms across countries and regions; mentoring and other professional development support activities; recognition of evaluation excellence and distinctive contributions to the discipline; promotion of multidisciplinary evaluation research, approaches and methods, including physical and natural sciences; and other (to be specified). Respondents were to indicate which functions they would support.

Most of the 451 respondents supported each listed function. The most highly supported functions were support for evaluation professionalism initiatives (76 per cent of respondents indicating it should be included) and harmonization of evaluation principles, guidelines and norms across countries and regions and support activities (73 per cent of respondents agreeing). Responses to the mentoring and advocacy roles were also strong. The lowest support (52 per cent of respondents) was for promotion of multidisciplinary evaluation research, approaches and methods.

A Strategic Challenge

Even with the survey findings showing strong support for an IEAc and the function of supporting evaluation professionalization within, some have still questioned the need for another potentially fragile evaluation organization further splitting funding for evaluation and possibly resulting in duplication of efforts. As discussed, IOCE has its hands full with building the capacity of VOPEs, and it does not provide an international voice on professionalization.

Also for consideration is IDEAS. Formed in 2002, IDEAS was conceived to help build an international community of development evaluators, and like most other evaluation societies, networks and associations, it has its own strategic priorities and agenda. IDEAS, again like evaluation organizations generally, has been underfunded and stretched to meet its own mandates, strategic objectives and workplans. Evaluation organizations are challenged in responding to national, regional or thematic connectivity needs. They have not been able to gather the resources needed to accelerate evaluation professionalization. They have only begun to focus on promoting systemic changes in the enabling environment.

Through its focus, global breadth of expertise, independence and stature, the proposed academy would support and complement the work of evaluation societies, networks and associations and partner closely with them. Its mandate would be to promote evaluation transformation, influence and professionalization. Its goal would be, among other things, to accelerate evaluation professionalization internationally. Although the goal

of evaluation is to create and disseminate knowledge for the public good, it cannot do so if it lacks influence.

Although evaluation has unique potential as a multi-discipline, a bridge across disciplines and a trans-discipline, it lacks the status, prestige and autonomy that other professions enjoy; it is not yet even listed in the classification of occupations that the governments of the world have adopted (ILO 2010). To climb towards the higher rungs of the occupation ladder, evaluation must overcome the following hurdles.

- High-quality evaluation education and training is scarce; evaluation has yet to be embraced as a mainstream academic discipline in most universities.
- Progress towards universal agreement on core (even using the 60 per cent rule already finding conceptual acceptance) guiding principles, ethical guidelines and competencies must be accelerated.
- The quality of evaluation work is highly variable, so commissioners frequently rely on other knowledge occupations to meet their needs.
- Most evaluation associations are too small to have a viable credentialing system.
- Young and emerging evaluators (YEEs) are not receiving the support they need to expand their evaluation knowledge and sharpen their skills.

To meet these challenges, the IEAc will undertake activities aimed at, for example, helping generate a larger supply of competent evaluators through targeted university advocacy and high-quality training programmes; securing a global consensus on core evaluator competencies through formal and mutual recognition; delivering brokering services to help commissioners identify competent evaluation practitioners; assisting evaluation associations with credentialing and increasing access to mentoring, peer reviews and tailor-made capacity-building initiatives aimed at YEEs.

As an academy that prioritizes evaluation transformation, it is imperative that the professionalization strand not be equated solely with credentialing. A primary line of effort for professionalization is identifying what the transformational evaluator looks like. How does that evaluator look and act differently from the traditional evaluator? What skill sets are implied? Should transformational evaluators seek to protect nature and advance human rights and responsibilities and have the skills and requirements to

do so? Should they aim for inclusiveness in evaluations by identifying and redressing the asymmetry of power relations embedded in evaluation practice? Should they commit to understanding and overcoming the drivers of violence and conflict, especially in evaluations of fragile countries and communities? These are the types of questions it is also important to address under professionalization in the transformation context. If there are to be transformational evaluations, we must have transformational evaluators.

The Enabling Environment

The sociology of professions demonstrates that finding a place in the sun in the professional arena is a highly competitive venture (Abbott 1988). Globally, all evaluation associations and networks have a combined membership of about 52,000 (and this includes double counting of members who belong to more than one association, often regional and national). This membership is scattered over about 140 associations, most of which by necessity are focused on national issues only. IDEAS is the only association with a global perspective, being founded to address problems without passports.

Total membership in evaluation associations is about one-fourth the membership of the Institute of Internal Auditors (200,000 members). Evaluation, still seen as the new kid on the block, is marginalized. The wider public is poorly informed about what evaluation stands for. Although some governments are adopting evaluation as a standard requirement – and although parliamentarians are increasingly drawn to evaluation – unlike accounting and auditing, evaluation is not close to being universally viewed as a standard statutory requirement for all interventions in the public, private and voluntary sectors.

Other disciplines (e.g. auditing, management consultancy, economics, econometrics, data science) encroach on the evaluation market. Some produce excellent evaluation work, and their contributions should be recognized, but as Dahler-Larsen (2013) has indicated, value-free social scientists who use the evaluation label to secure gainful employment should be challenged; evaluation is a vocation, and evaluations are public goods.

To protect and promote the evaluation brand, an IEAc would engage in public education and advocacy activities. It would help professionalize evaluation, fill the public information gap about evaluation and encourage routine use of independent evaluation for all social interventions across sectors and borders. Thus, it would be particularly active on the demand side of the evaluation profession.

Filling a Gap in the Global Evaluation Architecture

The challenge that the bracing vision of EvalAgenda implies is three-fold: the evaluation enterprise is far too small and fragmented to rise to current social and environmental challenges, most countries are not very far along the road to professionalization and the enabling environment is not supportive of evaluation. An IEAc would recognize and address these strategic priorities. To help address the crises of a troubled, interconnected world, such as climate change, biodiversity extinction and other problems without passport, the academy would have a global reach, and thematic networks would operate across borders. At the same time, it would be firmly grounded in local communities and civil society. Regional, national and local chapters would be sponsored and supported.

Thus, an academy would help evaluation become more 'international in the sense of being at the same time more Indigenous, more global and more trans-national' (Chelimsky and Shadish 1997). So that this process unfolds effectively, the academy would encourage knowledge communities to break out of their comfortable disciplinary silos; reach out to allied disciplines and close the gap between evaluation, social science theory, behavioural research and other knowledge occupations.

An example of a specific gap in the global evaluation architecture was recently provided in a personal communication, with support for an IEAc to address it. A blog posed the following question: Have you been looking for online evaluation courses but don't know where to start? Mikkolainen (2018) researched online courses and found many, but generally, they were costly, offered no guarantee of quality, did not generate a credible certificate and were often episodic or one-shot efforts. She therefore saw a role for an academy that would provide expert reviews and deliver quality assurance certificates to evaluation training courses.

Mission, Vision, Values and Guiding Principles

IEAc is being incorporated as a charitable organization in the United Kingdom. At this incipient stage, its focus has been on reaching a broad-based consensus about its mission, vision, values and guiding principles.

- *Mission.* The academy would act as a platform for innovation, creativity and collaboration in pursuit of evaluation's transformative impact, influence and professionalization.
- *Vision.* A world in which government, business and civil society are accountable, learn from evaluation and work together towards a healthy planet and societies that leave no one behind.
- *Values.* The academy would undertake to be self-reflective, open, respectful and honest in all its activities; strive for justice, diversity and tolerance and use the mnemonic ICCCI (integrity, compassion, courage, competence, inclusivity) to help keep its key values close.
- *Guiding Principles.* The academy would be responsive to the urgency of all the systemic changes needed to address the global crises endangering humanity and the health of the planet. It would seek to protect nature and advance human rights and responsibilities. It would respect the dignity and privacy of all evaluation stakeholders. It would aim for inclusiveness, in part by identifying and redressing the asymmetry of power relations embedded in evaluation practice. It would promote equity, gender equality, minorities and Indigenous peoples. It would commit to understanding and overcoming the drivers of violence and conflict, especially in fragile countries and communities. It would engage in transformational evaluations that make a difference. It would respect the subsidiarity principle in its work.

Organization

The IEAc is an inclusive, ambitious and, above all, activist venture existing primarily to identify, encourage and support evaluation initiatives geared to transformational change. It is a three-tier corporate organization consisting of a three-member interim organizing committee and a 14-member interim board of trustees that set policy, determine business priorities and oversee all academy activities; a 50-member council that manages programmes and projects in support of the IEAc's thematic directions and fellows who perform the work of the academy through projects responding to council-led, board-approved programmes. It is supported by a small secretariat.

The core assets of the academy are its fellows – and the relationships and partnerships that it will nurture with academia, societies, associations and networks. Fellows will include eminent evaluation thinkers, highly experienced evaluators, mid-career practitioners and YEEs who would be a

special target group. Fellows will be entitled to include the academy designation on their business cards and correspondence.

Individuals can apply for fellowships on the academy website (soon to go live), and a fellowship committee of the board will review applications and admit qualified applicants. Consistent with the academy's value of inclusivity and its rejection of elitism, the only criteria for fellows to be admitted to the academy are significant involvement in evaluation and related activities such as research, education, training, practice and management; commitment to academy purposes and values and willingness and ability to volunteer time and contribute to academy goals.

Initial Thematic Directions

The board has set the following initial thematic directions:

- *Evaluation research*: research on evaluation governance, management, models, methods and practices; science-based evaluation initiatives (e.g. complexity, systems thinking, mapping)
- *Evaluation advocacy*: promotion of sound evaluation policies, engagement with civil society groups, encouragement of corporate social responsibility-oriented evaluation
- *Evaluation education*: promotion of high-quality tertiary evaluation education in universities worldwide, especially in the Global South
- *Evaluation training*: incubation of innovative training projects focused on new evaluation frontiers, information technology, artificial intelligence, big data
- *Evaluation professionalization*: mutual recognition of guiding principles, professional ethics, competency frameworks, mentoring and professional development, support to YEEs' initiatives, prizes and awards
- *Transformational evaluation*: promotion of evaluations addressing transformational change, socially and environmentally sustainable development, Blue Marble evaluations
- *Evaluators without borders*: promotion of international evaluation exchanges, contributions to effective communications across evaluation cultures, amplification of Indigenous evaluation voices
- *Expert evaluation advice*: establishment of independent commissions, workshops, roundtables and expert meetings tasked with evaluation of critically important public policies and programmes; provision of independent, objective evaluation advice to improve

the validity, quality and social utility of evaluations in the public, private and voluntary sectors

The IEAc council is proposing and defining programmes for these thematic areas and asking the board for approval. The programmes will then direct all academy interventions (e.g. projects, events, commissions, transformational evaluations). Based on experience, some of these will be merged and others dropped or added. For example, evaluation education, training and professionalization might be merged into a single evaluation professionalization working group.

Business Model

Learned societies and professional academies can take many forms. Active engagement with a wide range of stakeholders will be required to design and fine tune the right organizational model and to design the operating procedures, but under any conceivable scenario, the energy and ideals of fellows will lie at the core of the initiative, and it is envisaged that the academy will evolve organically in response to fellows' aspirations and their local needs. Numerous iterations and course corrections are likely before the organization takes final shape. This is only the start of a consultative process.

To further its value of inclusivity, the academy will not rely on fellowship fees to sustain its work. Funding will be secured mainly through public and private donations for targeted initiatives and core academy expenses. Some funds could be raised from brokering and quality assurance fees. Costs also will be contained through fellows' contributions of their time (e.g. minimum of 10 days a year) and smart use of information and communications technologies.

The academy proposes to sponsor and oversee evaluations geared to transformational change and to set up independent commissions, workshops, roundtables and expert meetings tasked with evaluation of critically important public policies and programmes. Just like other academies, the IEAc will deliver remunerated expert evaluation services. Given its commitment to the *subsidiarity* principle in all its work⁶, the academy will not

⁶ Subsidiarity is a principle of social organization that holds that social and political issues should be addressed at the most immediate (or local) level that is consistent with their resolution. It is a general principle of European Union law (<https://en.wikipedia.org/wiki/Subsidiarity>).

compete with individual practitioners and private consulting firms that cater to the growing demand for *user-directed* evaluation services controlled by evaluation commissioners. It will steer clear of fee-dependent evaluations funded by decision makers.

This leaves wide scope for arm's-length funding from foundations, private individuals and the like for the provision of independent *evaluator-directed* activities. Such activities have generally gone by the wayside as user-directed evaluations have come to dominate the evaluation scene. In engaging in such activities, the academy may seek to collaborate with scientific academies, evaluation associations and societies.

The academy has secured funding for a small grant programme to support council members' and fellows' creative and path-breaking initiatives. Specifically, the small grants programme will make small (\$10,000 or less) financial contributions to volunteer-driven activities that hold promise for upscaling in line with board-approved strategies. Such grants would facilitate design of projects and events geared to transformational change and to the design, incubation and implementation of activities that support academy goals and priorities towards fulfilment of its charitable objects.

For example, grant requests would be entertained for fellows' preparatory work required for eliciting and organizing independent commissions, workshops, roundtables and expert meetings; pilot interventions in new and untested activity domains and catalytic work that may lay the foundation for innovative evaluation education, training and professionalization programmes and advocacy campaigns.

Grants might also facilitate advanced policy research work and sponsorship of high-quality academy publications. Grant applications will be subject to a peer review process managed by council members before submission to the Board Operations Committee (or a subcommittee) for approval. Proposed grant activities will be aligned with the thematic strategies of the academy and its values and principles. Due diligence processes will be put in place before the small grants programme is formally launched.

Conclusion

The IEAc has been proposed in part to accelerate evaluation professionalization, increase evaluation influence and transform evaluation practice. A brief survey that IDEAS conducted on the concept of the academy found strong support for it and a focus, among other things, on professionalization. Although the partnership process is just underway, the mission,

strategy, guiding principles, values, organization, thematic strategies and business model of the proposed academy have been sketched out. The strategies include a strong focus on evaluation professionalization. The academy has a working board and a large, active council that is translating the thematic strategy into programmes. Incorporation as a charitable organization is well underway, and its website will soon be live and ready to take fellowship applications. It has a strong mandate and an opportunity to transform evaluation professionalization.

References

- Abbott, A. 1988. *The system of professions: An essay on the division of expert labor*. Chicago: University of Chicago Press.
- Annan, K. 2009. Problems without passports. *Foreign Policy*. <https://tinyurl.com/49dpxz57>
- Baptiste, L., and V. Iese. 2019. The transformative agenda for evaluation in small island developing states: The Caribbean and the Pacific Island. In: *Evaluation for transformative change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R. D. van den Berg, C. Magro and S. Salinas Mulder, 71–88. <https://ideas-global.org/transformational-evaluation/>
- Chelimsky, E., and W. Shadish (eds.). 1997. *Evaluation for the 21st century*. Thousand Oaks, CA: Sage Publications.
- Dahler-Larsen, P. 2013. *The evaluation society*. Stanford, CA: Stanford University Press.
- Gauthier, B. 2019. The professionalization of evaluation practice: definitions, process, and trade-offs. Keynote address at the Conference of the Swiss Evaluation Society. <https://tinyurl.com/hjs9je6k>
- Gounou, A., and G. Perez Yarahuan. Transforming national M&E systems to serve Agenda 2030: Africa and Latin America. In: *Evaluation for transformative change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R. D. van den Berg, C. Magro and S. Salinas Mulder, 51–70. <https://ideas-global.org/transformational-evaluation/>
- Heider, C. 2015. Professionalizing evaluation: What more is there to say? Blog post. <https://tinyurl.com/vzztwbmk>
- Heider, C. 2018. Professionalizing evaluation. World Bank blog. 28 April. <https://tinyurl.com/78wjzhz9>
- ILO (International Labour Organization). 2012. *International Standard Classification of Occupations: ISCO-08*. Geneva: ILO.
- Jacob, S., and Y. Boisvert. 2010. To be or not to be a profession: Pros, cons and challenges for evaluation. *Evaluation* 16 (4): 349–369.
- Julnes, G., and K. Newcomer. 2018. Supporting the AEA mission of improving evaluation practice, methods, and use: The timely value of a national academy of evaluation. *American Journal of Evaluation* 39 (4): 586–595.

- Mikkolainen, P. 2018. Interested in learning about evaluation online? Start here. Blog post. <https://tinyurl.com/rpwje5ps>
- Morra Imas, L. 2017. Professionalizing evaluation: A golden opportunity. In: *Evaluation of Agenda 2030: Providing evidence on progress and sustainability*, (eds.) Rob D. van den Berg, Indran Naidoo, Susan D. Tamondong. 65–80. Exeter, UK: IDEAS. <https://tinyurl.com/txh66ufc>
- Morra Imas, L. 2010. Professionalizing development evaluators. Travelling the road to results. In: *From policies to results: Developing capacities for country monitoring and evaluation systems*, (ed.) M. Segone, 171–188. New York: UNICEF.
- Morra Imas, L. 2019. IDEAS Concept Note on IEAc. 8 April. Unpublished. Available from ieac.coordinator@gmail.com.
- Ofir, Z. 2020. Transforming evaluations and COVID-19, Part 4. Accelerating change in practice. Blog post. <https://tinyurl.com/aw5x2xnt>
- Patton, M. Q. 2020. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Picciotto, R. 2011. The logic of evaluation professionalism. *Evaluation* 17 (2): 165–180.
- Quesnel, J. 2010. The professionalization of evaluation. In: *From policies to results: Developing capacities for country monitoring and evaluation systems*, (ed.) M. Segone, 164–170. New York: UNICEF.
- Schwandt, T. 2017. Professionalization, ethics, and fidelity to an evaluation ethos. *American Journal of Evaluation* 38 (4): 546–553.
- UNEG (United Nations Evaluation Group). 2016. Professionalization of evaluation concept paper. <https://www.vopetoolkit.ioce.net/es/node/582>
- Van den Berg, R.D., C. Magro and S. Salinas Mulder (eds.). 2019. *Evaluation for transformative change: Opportunities and challenges for the sustainable development goals*. <https://tinyurl.com/2awd52xs>



CHAPTER 11

Evaluation for Transformation: What Will It Take?

ROBERT PICCIOTTO

Abstract. Given recurrent health emergencies, rapid environmental degradation, pervasive insecurities and the rising popular anger that the unmet promises of modernity in liberal and authoritarian regimes alike have triggered, populism is on the rise, the knowledge professions are threatened and social transformation is imperative. Thus, evaluation faces its own transformation challenge. New policy directions will be required to transform the enabling environment of evaluation practice. Specifically, the process that the neoliberal and evidence-based waves of evaluation diffusion induced, which transformed evaluation into a private good, must be reversed. For evaluation to restore its public good character, it must break the chains of the market-based governance model currently in place and increase its influence by moving up the occupational ladder. This implies acquiring all the interrelated features of professionalism: an ethical charter, expert knowledge, proven competencies and self-management. There is no shortcut.

Introduction

Transformation first emerged as a fashionable buzzword in a corporate world beset by rapid change (Bucy, Hall and Yakola 2016). Fuelled by the climate change crisis and the COVID-19 pandemic, it is now a universal preoccupation of policymakers confronted by recurrent health emergencies, rapid environmental degradation, pervasive international criminality and unprecedented financial volatility.

Evaluators are striving to rise to the occasion. They recognize that transformed evaluation methods and mindsets are required. Even before the COVID-19 crisis hit, evaluation conferences, publications and blogs had highlighted the need for new thinking and new methods, but the evaluation community has yet to recognize the full implications of the transformation challenge. Thus, its professionalization initiatives have been modest, scattered and uncoordinated¹.

The chapter is in five sections:

- First, it takes stock of the human progress associated with past development transformations, as well as the risks to humanity associated with current public policies.
- Second, it probes the widespread discontent found within and outside the evaluation community.
- Third, it identifies gaps in the evaluation community consensus regarding what must be done.
- Fourth, it demonstrates that four specific interrelated professionalization challenges must be met for evaluation to be transformative.
- Fifth, it offers concluding remarks.

Past Transformations: Achievements and Drawbacks

The imperative of a basic reorientation in policy directions is not unprecedented. A major transformation challenge faced policymakers 75 years ago when the international development mission emerged out of the ashes of World War II. For the first time in world history, official and public opinion in Western societies acknowledged the need to attend to the 'urgent problems

¹ The International Organization for Cooperation in Evaluation has deliberately limited its role as a neutral source of information on professionalization actions undertaken at national and regional levels, which have been few and far between.

of economic development of underdeveloped countries' (United Nations 1949, 251). Thus, the victorious allies adopted international aid as a major foreign policy instrument, and economic development became the primary public policy objective of the governments of less-developed countries.

At the same time, evaluation, the 'new kid on the block' of the academy, came into existence. It embraced an inspiring mission: connecting the then highly esteemed social sciences with government policymaking for the good of all (Shadish and Luellen 2005). This was a time of optimism, faith in government and belief in international cooperation. In relatively short order, a handful of far-sighted evaluation pioneers assembled a toolkit of evaluative methods, launched specialized publications and created evaluation associations.

The Great Convergence

A major transformation of the world economy ensued. The admixture of growth-oriented policies, official development aid and foreign direct investment was providential for many developing countries. Although it left more than 1 billion people in 47 least-developed countries behind, it induced a gradual convergence in average per capita incomes between the North and the South. Lifted by increased prosperity in emerging market economies, average global per capita incomes increased from \$3,300 in 1950 to \$18,000 in 2019 (Roser 2019; Statistics Times 2021).

By the 1990s, driven by the dynamic economies of Asia, the developing world began to become the engine of the global economy. As a result, despite sharply increased within-country inequalities, global inequality declined (Bourguignon 2016), as did the share of the absolute poor of the world population (Revenga and Dooley 2019). Whereas some 60 per cent of the world population lived in extreme poverty in 1950, only 10 per cent did before the COVID-19 pandemic struck (Roser and Ortiz-Ospina 2013).

Social indicators improved as well. By 2019, global average life expectancy (73 years) reached higher levels than in any country in 1950 (Roser, Ortiz-Ospina and Ritchie 2013). The prevalence of chronic undernourishment and catastrophic famines also declined with the advent and dissemination of new agricultural technologies (Roser and Ritchie 2019). Literacy, especially female literacy, became widespread².

² The gender gap in literacy is declining for all regions, with Europe and Central Asia, Latin America and the Caribbean, and East Asia and the Pacific having nearly closed this gap, although other regions are far behind. More information available at Wadhwa 2019.

Other markers of human security and welfare followed suit. The global number of battle deaths fell dramatically (Roser 2016), as did homicide (Roser and Ritchie 2013) and genocide deaths (Roser and Nagdy 2013). Improvement in global well-being indicators have been variously ascribed to Enlightenment values, scientific progress and international cooperation (Pinker 2018). Irrespective of their antecedents, these remarkable achievements have come with a heavy price tag for society and the natural environment.

Social Costs

While nearly half of the world is still striving to subsist on \$5.50 a day or less, the world's richest 1 per cent have secured twice as much wealth as close to 90 per cent of the world population. Social cohesion has been sorely tested under the sway of extraordinary increases in inequality. With globalization in full swing, the world witnessed a massive reallocation of labour-intensive work towards emerging market economies, combined with de-industrialization in developed nations.

In Western countries, working-class earnings eroded because of runaway outsourcing to low-wage countries and automation. In Australia, Canada, Ireland, the United Kingdom and the United States, mortality rose due to drug overdose, alcohol and suicide (Case and Deaton 2020). This epidemic of distress spared most college-educated citizens so that unequal access to higher education aggravated the social costs of income and wealth inequalities.

A Planet at Risk

The 2008 financial crisis demonstrated the huge risks to livelihoods associated with financial globalization. Even more insidious and deadly, the silent climate change crisis will inflict escalating damage on societies and the natural environment (Patton 2020). Given persistent international reliance on carbon-intensive economic policies, the world is heating up; it is as if humanity is facing the fate of proverbial frogs immersed in water gradually reaching the boiling point.

The fossil fuel-based energy dependence of the world economy is mostly driving climate change, but human diets are also to blame. Animal agriculture, increasingly dependent on cruel and polluting factory farming, is responsible for 13 per cent to 18 per cent of human-caused greenhouse gas emissions. Furthermore, the relentless growth of the livestock economy

worldwide has been detrimental to human health because red and processed meats increase the risk of cancer. The stark reality is that current policies and consumption patterns are not sustainable. Four to five planets would be needed to accommodate all countries at current Western per capita income levels.

Public Discontent and Its Consequences

Unsurprisingly, huge public demonstrations have called on governments to step up the actions that are patently and urgently needed to restrict the rise in global temperatures to 1.5°C, the modest goal agreed under the 2015 Paris Agreement. Students, workers and professionals have joined climate strikes. Popular protests, often led by youth movements, have spread in reaction to air pollution, plastic waste and rising sea levels.

The Davos policy consensus of the rich and famous has long praised disruptive innovation and favoured marginal feel-good policy improvements. This narrative is no longer persuasive (Guardian 2019). Global opinion surveys confirm that distrust of elites and dissatisfaction with the workings of liberal democracies are rife. The growing popular discontent is tied to economic grievances, politicians' unresponsiveness to popular demands and widespread suspicion of out-of-touch intellectuals (Wike, Silver and Castillo 2019). As a result, populism has spread, distrust of science has risen and the public is turning to authoritarian leaders.

Democracy in Decline

Whereas in the 1990s, after the implosion of the Soviet Union, democracy was on the march, the prevalence of liberal democratic regimes began to fall in 2005, just as it had in the 1930s. The latest aggregate Democracy Index that the Economist (2019) Intelligence Unit compiled was the lowest recorded since the index was first published in 2006.

In 2019, only 22 countries, home to 430 million people, were full democracies, whereas more than one-third of the world's population lived under strict authoritarian rule. In Russia, Turkey, Hungary, Poland, the Philippines, Brazil, India and the United States (which was downgraded from the full democracy to the flawed democracy category in 2016), authoritarian leaders rose to power. In sub-Saharan Africa, 23 countries saw their democracy scores decline, whereas only 11 improved.

The Policy Context

Given the stubborn facts described above, it is high time for a fundamental transformation of policy frameworks. They should be directed away from runaway, inequitable, unsustainable growth towards enhanced human security. This requires diligent precautionary policies that hedge for the worst in the face of rare but potentially catastrophic systemic risks (Taleb 2007). It also means a dominant role for the state and competent administration. Unfortunately, in many countries, governments have taken a back seat and let business interests shape policies characterized as socialism for the rich and capitalism for the poor.

It is time to confront the powerful lobbies that have captured policy-making. Powerful, profit-driven, unregulated, digitalized companies have undermined social cohesion. The social media they have created act as echo chambers for fact-free opinion and extremist politics. Taxation of the super-rich has declined, individual and corporate tax dodging have become widespread, regressive value-added taxation that penalizes the poor has risen, budget austerity has constrained social spending and global monopolies have escaped regulation.

What Kind of Transformation?

Meeting the new transformation challenge will not take place absent an understanding of what has happened. The hard-won lessons of experience must be drawn. Given that the fruits of innovation and growth have not been equitably shared and that the future of the planet is threatened, the New Public Management ideas that have shaped policy should be discarded (Ventriss 2000). Yet the rich and powerful remain committed to rigid neoliberal doctrines and will seek to exploit the disorientation that the COVID-19 catastrophe has caused to secure adoption of radically conservative policy measures.

This would be a repeat of their past behaviour, for example when they exploited the fears that the 9/11 attack elicited to erode civil liberties; allowed real estate developers to displace thousands of poor households to make room for luxury hotels and apartments following the ravages of the Indian Ocean Tsunami, the Haiti earthquake and Hurricane Katrina and took advantage of debt crises to push through economic shock therapies and budget austerity policies (Klein 2007).

Thus, the shape of the coming social transformation is in doubt. Although capitalism has captured the commanding heights of the global economic system, it has taken different forms in the liberal meritocratic

states of Western industrialized countries and political capitalist states such as China (Milanovic 2019). It is far from clear which of these two capitalist typologies will prevail, because in both cases, the economic and political elites have come together with predictable results: growing inequality and corruption that could in time trigger a popular backlash.

The Anatomy of Disenchantment

What explains the sharply contrasting policy beliefs prevalent in contemporary society? According to Max Weber (2002), modernity arose following the erosion of religious faith and magical thinking as soon as the Enlightenment demystified contemplation and rejected the sacramental mediation of salvation. Suddenly, rationality overcame superstition, work became valued for its own sake and profit making that had previously been despised as a manifestation of greed and avarice became honourable. Science acquired prestige, but its success in disenchanting the world has been neither complete nor definitive, because reason without faith is alienating, and science cannot answer questions about values and morals. Magical thinking, rejection of scientific findings and distrust of expertise are making a comeback. All knowledge occupations, including evaluation, have become vulnerable to the populist backlash.

In the wake of disenchantment, a decisive shift of focus from faith-based tradition to rational action gradually transformed everyday life and laid the foundations for the triumph of capitalism. Social action, driven by custom and routine, and affective social action, driven by instinct, impulse, anxiety or desire, were hemmed in. Reason displaced emotion and made room for scientific inquiry, private enterprise and democratic debate. This paved the way for the march of modernity and the ascent of liberal democracies, which in turn triggered a revolution of popular aspirations fuelled by a universal culture of individualism, egalitarian ambition and deep longings for the wealth and status that less than 1 per cent of the world population enjoyed.

The huge chasm that divides the elite from the masses has opened a space that demagogues, cultural supremacists and brutal extremists now occupy. The new hypercapitalist, neoliberal world has encouraged 'the suspicion – potentially lethal among the hundreds of millions of people condemned to superfluosity – that the present order, democratic or authoritarian, is built upon force and fraud; they incite a broader and more apocalyptic mood than we have witnessed before. They also underscore the need for some truly transformative thinking about the self and the world' (Mishra 2018, 346).

Evaluation Rationalities

The intellectual legitimacy of evaluation is grounded in theoretical rationality, whereas its social legitimacy hinges on the other kinds of rationality it adheres to as a guide to action. When evaluation is institutionalized, formal rationality dominates so that power captures evaluation. When practical rationality has the upper hand, powerful interests induce self-serving evaluative biases. Only substantive rationality at the service of instrumentally rational and ethical evaluation is socially legitimate.

It follows that the same secular forces that transformed the world have shaped the trajectory of the evaluation occupation. At the outset, evaluation pioneers strove to inject rationality into the untidy world of politics. Thus, Donald T. Campbell, a scientist, humanist and generalist, joined the evaluation ranks in the mid-1960s when colleagues induced him to champion quantitative methods in social research (Picciotto 2019a). The ideal 'experimenting society' shaped by the *experimental wave* of evaluation diffusion was an open plea to policymakers to subject social programmes to systematic, quantitatively rigorous evaluation (Campbell 1971).

Although this technocratic vision was compelling, the *dialogic wave* of evaluation diffusion, an even more inspiring model of inclusive, value-driven evaluation, replaced it in the 1970s. These were the halcyon days of evaluation as a force of good, but the faith in evaluation that it elicited evaporated when this exceptionally innovative period of evaluation history came to an end when a *neoliberal wave* engulfed the evaluation discipline in the 1980s.

This is when market thinking infiltrated all sectors of society, and value-free management consultants working across borders were recruited to serve decision makers subservient to free-wheeling capitalist interests. The same ethos is sustaining the *evidence-based wave* that characterizes evaluation diffusion today. The evaluators that surf this wave do not challenge neoliberal ideas. They advocate a renaissance of scientific experimentation while stressing accountability, value for money and customer satisfaction under the slogan 'what matters is what works!' (Vedung 2010).

In parallel and paradoxically, disenchantment with evaluation spread just as evaluation crossed borders. By now, propelled by the development cooperation industry, evaluation practice has become genuinely 'international in the sense of being at the same time more indigenous, more global and more transnational' (Chelimsky and Shadish 1997, xi).

Evaluation Disenchantment

Transformational evaluation, which Donna Mertens (2005, 422–423) defines as evaluation that pursues ‘the goal of bringing society to a point of greater equity and justice’, has captured the imagination of evaluation practitioners, but the mood of influential evaluation thinkers has nevertheless darkened.

As the neoliberal and evidence-based waves washed over evaluation practice, belief in the public interest mission of the evaluation discipline faltered, and eminent evaluation thinkers began to ask the same questions about evaluation that they have routinely asked of others: Does evaluation ‘work’? What has been achieved and at what cost? The results of this introspection have been sobering.

Thus, Peter Dahler-Larsen (2012, 231) has deplored the high transaction costs and the detrimental effects of linear evaluative thinking on creativity and innovation: ‘it is time to consider...whether the marginal utility of evaluation may be decreasing and whether there are sometimes good reasons for evaluation fatigue’. He has also observed that ‘in recent years, we have witnessed a boom in evaluation... It is as if there is no limit to the feedback loops...as if the insatiable evaluation monster demands more food every day’ (Dahler-Larsen 2012, 1).

By now, goal achievement evaluation has become integrated as an administrative routine in all sectors of the world economy. Co-opted, institutionalized and routinized, evaluation is now shaped by buyers’ preferences, the range of evaluation questions has become more restricted and manager-oriented evaluations tend to neglect the public interest. These shifts in orientation of evaluation practice have blurred the boundaries between evaluation and other knowledge occupations.

Specifically, evaluation has become conflated with auditing, inspection and other means of social control that the public perceives as costly, ritualistic and disruptive. The illusory comfort offered to managers facing the realities of an uncertain and turbulent operating environment has sustained intrusive oversight, detailed record keeping, intense bureaucratic scrutiny, constant pressure to demonstrate rapid results and mandatory use of simplistic performance measures.

To be sure, not all evaluation thinkers are resigned to the current submissive status of the evaluation occupation. They remain wedded to the lofty evaluation ideals present at the creation of the discipline. For example, Robert Stake (2016) has consistently asserted that evaluation can serve the pursuit of equity; Karen Kirkhart (2015) has advocated advancing social

equity through evaluation and Jennifer Greene (2012) has encouraged evaluators to commit to equity-minded, critical evaluative habits.

Nevertheless, these professions of faith in evaluation coexist with gnawing self-doubt because evaluation, once a public good, is now bought and sold in a market where evaluators frequently yield control over their work to commissioners beholden to power holders. By now, evaluation is widely perceived as an enterprise that evaluators, decision makers, and evaluation commissioners jointly own; that is, the substantive rationality on which the ethical foundation of evaluation rests has given way to the practical and formal rationalities that the rich and the powerful have mobilized.

Gradually, evaluation models that emphasize social justice, democracy and inclusivity have lost ground to a usage-focused culture geared to the achievement of managerial goals. All too often, evaluation is relegated to the fulfilment of managers' needs for data gathering, and the biased computer algorithms of the new information economy have escaped evaluation while assuming enormous influence as standard management instruments wielded by value-free data scientists.

In parallel, evaluators have been pushed towards the periphery of policymaking. Thus, evaluation is now widely conceived as an enterprise rather than a vocation, a business venture rather than a special calling, 'one tool among many for the improvement of policies, learning and social change' (Furubo and Stame 2019, xv) These trends must be reversed for evaluation to meet the huge transformation tasks implied by the diagnostic offered in the previous section.

Gaps in the Evaluation Community Consensus

Thankfully, the evaluation community has seized on transformation as an objective worth striving for. The transformation vision is akin to a flickering light at the end of a dark tunnel. It implies radical changes in evaluators' mindsets towards risk management, conscious citizenship and planetary well-being assessment (Ofir 2019b). It draws on transformation theory that knits together the complexity sciences, systems thinking, feedback loops and network effects.

The need for refinement in evaluation competency frameworks to address the challenges of sustainable development and the digital revolution more effectively is also acknowledged (Ofir 2019a). In parallel, the mainstream evaluation community has focused on shifting the focus of the evaluation enterprise from assessments of individual interventions to

systemic policy changes, but with few exceptions (Mathison 2016), the current consensus has not reckoned with the broader implications for evaluation of the demise of democratic capitalism, the surge of neoliberalist ideas, the manufacture of popular consent by a captured elite or the rising influence of business interests on political decisions.

Although these transformation ideas are injecting fresh energy and renewed hope within the evaluation community, most evaluation thinkers, managers and practitioners have yet to come to terms with the extent to which the enabling environment for evaluation has become inimical to democratic, culturally sensitive, transformative, independent evaluation (Picciotto 2015).

In the contemporary evaluation scene, results-oriented evaluation clients who hold the purse strings and impose tight constraints on evaluation practice have set aside progressive evaluation models. Asking evaluators to break away from the intellectual straitjackets imposed by commissioners is a tall order because evaluation is now a commodity subject to market forces (Mathison 2016).

Evaluation community leaders have focused on the supply side of the evaluation market, yet power holders dominate the neglected demand side in a market in which economists, auditors and management consultants have the upper hand and independent evaluators committed to social and environmental justice must summon extraordinary courage and strike painful compromises as they struggle to secure contracts in an increasingly competitive market.

How then can the evaluation community overcome the pervasive conflicts of interest that mar evaluation practice in pharmaceutical evaluation, education evaluation and financial evaluation, as Ernest House (2016) conclusively documented? How can client-centred, utilization-focused (Patton 2008) or developmental evaluations that Michael Quinn Patton (2010) has brilliantly advocated for embrace empathy, cultural sensitivity, inclusiveness, self-awareness and reflexivity within the constraints imposed by having 'skin in a game' that vested interests control?

Because evaluators cannot readily bite the hands that feed them while making a living, how realistic is Thomas Schwandt's (2019) vision of a post-normal evaluation future grounded in ethical accountability, co-production, practical reasoning and support for citizen engagement in democratic decision-making processes? Is his caricatural depiction, offered in jest, of evaluators' submissive roles as scientific watch dogs, policy guide dogs or subservient lapdogs closer to the uncomfortable contemporary

reality? What then is to be done to ensure that those who pay the piper do not call the evaluation tune (Datta 2016)?

This is where the promise of professionalization comes in. Talcott Parsons (1968) visualized an ideal world in which the professions would acquire enough power to minimize the tyranny of the state and the excesses of capitalistic exploitation. This is the direction of travel that the evaluation community should adopt if, beyond its confirmed status as a legitimate and distinct discipline³, it opts to pull itself up by the bootstraps to play a more influential and effective role in society.

What Will It Take?

Vision without reality is hallucination. It should be clear by now that evaluation must be fundamentally transformed to help transform society and address the interrelated, systemic policy dysfunctions at the local, national and global levels that have led humans to deplete the planet's resources, aggravate inequalities and increase the vulnerability of society to recurrent, catastrophic disruptions in lives, jobs and social cohesion.

Facing Reality

Evaluation is still an infant industry. Globally, all evaluation associations and networks that EvalPartners surveyed have a combined membership of 32,000, and this includes double counting of members who belong to more than one association. This is less than one-fifth of the membership of a single association of internal auditors (e.g. the Institute of Internal Auditors has 175,000 members). There are some 1.2 million accountants and auditors employed in the United States alone. Growth in evaluation practice will require market diversification well beyond the government sector.

The market-led revolution that has swept over the global system since the turn of the century also means that the evaluation community should expand its scrutiny of the private sector and the growing philanthropic sector. It should target social impact funding initiatives, Big Data algorithms, and non-governmental organizations. It should build stronger relationships with community-level organizations and advocacy groups. It

³ Many social researchers reject evaluation's claim to the status of an autonomous discipline, let alone a profession. They do not think that it is sufficiently systematic, coherent or theory driven – 'a helter-skelter, mishmash, a stew of hit-or-miss procedures' (Davidson 2005).

should master the new information technologies and find cost-effective ways to deliver adequate, timely evidence to decision makers.

This implies a new way of doing business and vigorous evaluation advocacy. The evaluation brand should be more sharply defined and better protected. Currently, the wider public is poorly informed about what evaluation stands for. Evaluators are regularly confused with auditors and social researchers. High-quality tertiary evaluation education is scarce⁴. The discipline has yet to reach universal agreement on guiding principles, ethical guidelines and competencies for evaluators.

Evaluators do not control access to the evaluation discipline, and as a result, the quality of evaluation work is highly variable. Anyone can pose as an evaluator. It is high time for the evaluation world to face up to these realities. To be sure, evaluators have recognized the need to adapt their methods, refine their competency frameworks and raise their sights from individual interventions to systematic examination of higher-plane systems and policies. They have also sought to build evaluation capacities and develop training opportunities.

The Limits of the Current Consensus

Evaluation community leaders have neglected to recognize the full implications of the transformation challenge for their practice: the need to enhance the influence of evaluation as a specialized, autonomous knowledge occupation. Although it evokes special privileges, a professional label would add prestige and enhance the influence of evaluation practitioners in society. As a result, the enabling environment would be transformed so that it recognizes evaluation as a profession, ensures that all sectors of society understand and appreciate the value of evaluation and explicitly includes evaluation through national policies and other governance and regulatory instruments.

These changes on the demand side of the market were among the Global Evaluation Agenda (GEA) goals that the evaluation community adopted at a historic meeting held in Kathmandu (Nepal) in 2015 – the Year of Evaluation (EvalPartners n.d.). Although professionalization has become a central concern of communities of practice in evaluation, no consensus

⁴ Although there are 25,000 universities in the world, the American Evaluation Association (AEA) has identified only 80 university evaluation programmes worldwide.

has materialized as to what this means for the evaluation community, given persistent concerns about the dark side of professionalization.

Evaluators remain ambivalent and divided regarding the desirability of systematic professionalization⁵, and evaluation associations have sorely neglected the advocacy dimension of their remit.

Scepticism is certainly warranted. All professions must guard against unreasonable barriers to entry, elitism, self-serving practices, bureaucratization and ossification. Evaluation is not immune to the risks that monopolistic practices, professional self-interest and a narrow focus on methodologies and technicalities pose to innovation, creativity and intellectual openness. The restrictive characteristics of professionalization should be resisted, and paths towards evaluation excellence should be laid for young, emerging evaluators.

This said, without full-scale professionalization, evaluation will continue to be marginalized, evaluation disenchantment is likely to persist and the gap between lofty aspirations and results will grow. Fortunately, the basic elements of a sound professionalization strategy are at hand; an authoritative 'sociology of the professions' literature has identified four main drivers of professionalization. They are summarized here.

Adopt a Professional Ethos

First, the determination to work in the public interest is a prerequisite for securing the franchise to operate with autonomy in the public sphere. For example, such professions as medicine have adopted charters that champion the primacy of patient welfare and the promotion of social justice in the health industry.

In a powerful essay, Thomas Schwandt (2017) has deplored the glaring absence of vigorous discussion of what evaluation aims to add to society and the social good it seeks to serve. This must be done for a *professional ethos* to be defined and enforced. As things stand, evaluation simply means 'the process of determining merit, worth, or significance' (Scriven 2007, 1). This widely accepted definition is elastic and permissive. It implies that evaluation may focus on one, two or all three of these evaluative dimensions.

This ambiguous remit has allowed compliance audits (merit) and consumer guides (worth) to masquerade as evaluations. It has failed to put values and the public interest at the core of what evaluation should be in

⁵ Ian C. Davies captured the collective progress towards professionalization, as well as the widespread doubts still prevalent in the evaluation community in his preface to Davies and Brummer (2015).

the age of transformation. Missing from the widely accepted evaluation definition is the moral obligation for evaluators to promote the public good and take an ethical stance that they can defend with scientifically acceptable evidence (Scriven 2016).

Adopting a more demanding professional ethos would also tighten boundaries around evaluative inquiry. It would no longer allow exclusion of basic moral concerns. The ethical guidelines that evaluation associations issue would have to be expanded and improved. They currently concentrate on individual evaluators and neglect the evaluands. Newly upgraded ethical standards would address the responsibilities of evaluation commissioners. They would make clear that merit assessments include ethical evaluation of social interventions goals, that worth assessments comply with progressive values and that significance assessments would focus on the public interest.

Thus, new forms of formal and practical rationality reflecting the common good would be mobilized to buttress the substantive rationality of the evaluation discipline. Once evaluators take control of their own work and commit to a common professional ethos, all evaluations would make adequate room for expert estimation of the indirect and unintended social and environmental effects and would address the ethics of all evaluation participants. Evaluators would be enjoined to refuse evaluation assignments intended as subterfuge (e.g. evaluations commissioned to delay needed action, to duck responsibility, for window dressing or for public relations). They would subject evaluation terms of reference to critical review.

Upgrade Expertise

Abundant evidence has been adduced to demonstrate that modern economies require specialization, which in turn explains the ascent of professions. Thus, selected individuals who have undergone specialized education at the tertiary level followed by substantial exposure to skilled practice and periodic updating of their expert knowledge have come to perform sensitive and complex tasks that affect human welfare and facilitate the smooth functioning of society.

Meeting the accountability and learning requirements of society, especially in a context of rapid transformation, is a specialized task. Evaluation is not an amateur sport. Evaluators are not mere technicians. They cannot aspire to handle the multidisciplinary dimensions of their practice without a solid general education at the tertiary level. Furthermore, they need to have a firm grasp of specialized evaluation methods, appreciate their potentials and understand their limitations.

Although experience and practice are critically important, formal education is a pillar of professionalism. Evaluation work defies standardization. Evaluators are not only craftsmen and technicians; they also have privileged access to relevant and valid theories and a capacity to mobilize expertise from a wide range of disciplines. They are equipped with up-to-date knowledge, specialized skills and sound judgment.

Hence, evaluation professionalism cannot be divorced from the institutionalization of occupational expertise through high-quality tertiary education. Unfortunately, evaluation has yet to find its rightful place in the academic sun. Evaluation departments in universities are a rarity. Few universities have recognized that evaluation is not social research and that the standard curriculum should make room for evaluation.

Control Access to the Practice

Controlled access to the practice is a defining feature of all jurisdictional contexts within which professions operate in modern societies, although the potential restrictions on entry can be modulated over a wide range. Different models have been adopted in various combinations within diverse country contexts to select members fit to join the professional cadres of meritocratic governance systems.

A person qualified to perform a job or task earns designation after validation by a professional body acting to safeguard the reputation of a discipline. Credentialing confirms proof of completion of specified training and experience. Certification uses a variety of instruments to confirm possession of the basic knowledge, skills and experience required to perform professional work. Licensing implies legal control over the ability to practice, including the power to remove the license if professional standards are not being adhered to.

Thus, the enabling environment for professions is not monolithic. It depends on the national administrative and legal context and the leadership orientation of the sponsoring occupational group. Governments, professional associations, accredited academic establishments or a combination of these can exercise control over entry, but such screening is needed to enhance service quality and facilitate consumers' choice of service providers by managing the risks that users face when they select professional experts through personal contacts, word-of-mouth testimonials or trial employment. Designation, credentialing and licensing systems are all designed to reduce transaction costs in the evaluation market and limit prohibitive malpractice risks.

Ensure Professional Autonomy

Professionalization scholars define professionalism as the existence of set of institutions that allow members of an occupational group to make a living while controlling their own work (Freidson 2001). The principles that animate professions differ fundamentally from those of competitive markets and of public or private bureaucracies. In Eliot Freidson's words (2001, 221), members of a profession even 'claim the right to judge the demands of employers or patrons and the laws of the state, and to criticize or refuse to obey them'.

In its ideal form, it is the profession itself that directs all aspects of its governance through such things as controls on recruitment, quality of training, approval of professional guidelines and enforcement of ethical standards. This allows regulation of the supply of professional services, prices and fees. All contemporary models of professionalism stress the importance of self- management and autonomous control over occupational practices.

The need for autonomous self-management became self-evident once evaluation was commodified. This implies setting administrative rules, conducting peer reviews of work quality, disciplining members and in extreme cases stripping them of their designation. These measures can be considered monopolistic and self-serving, but the need to minimize the market disorder that inevitably prevails when quacks and amateurs can enter the fray with impunity, thus triggering distrust, confusion and poor service quality, amply justify them.

Concluding Remarks

What then are the takeaways from this chapter? First, the retrospective it put forward highlighted the economic gains, as well as the severe social shortfalls and existential risks, associated with past economic transformations. To be sure, transformation is an overused term, but it captures the fresh receptivity to change that widespread public dissatisfaction with the state of the world triggers. The dominant free market thinking that has swept over society since the mid-1980s delivered economic growth, but it shaped socially and environmentally unsustainable outcomes. The resulting disenchantment that sociologists presaged did not spare evaluation.

A Limited Consensus

The evaluation community has recognized the transformation challenge. It has begun to refurbish its assessment criteria to address transformational change. It is seeking to be more relevant, timely and technology savvy. It is tightening its competency frameworks. It is shifting its focus from individual interventions to the higher plane of policy. It is adding to the evaluation toolkit by drawing on systems thinking and connecting to the complexity sciences. It is exploring the premises and principles that should govern evaluation in the Anthropocene Age. These are much-needed supply-side reforms, but the demand side of evaluation practice also needs to be transformed.

Ethics Matter

Transcendent values add moral substance to the technical content of any discipline. Recapturing the ideals that evaluation evinced in its formative years is key to its future, but to do so, evaluators must face reality; evaluation has been commodified and captured, and it needs to break free from the shackles of power holders and the tyranny of market forces. A dominant role for ethics would give meaning to and justify evaluation independence and self-management. Hence, formal attachment to a distinctive, demanding professional ethos is a fundamental prerequisite of evaluation transformation. All evaluations should be progressive.

Evaluators will be able to distinguish themselves from auditors, management consultants and social researchers primarily by putting their social conscience to work. This implies a tighter definition of what evaluation is – as well as upgrading of evaluation principles and guidelines. Putting the public interest at the core of evaluation practice would contribute to the effectiveness of evaluation advocacy and facilitate brand differentiation. In turn, this would help nurture a latent public demand for principled, no-holds-barred evaluation. It would help evaluation grow.

Knowledge Is Key

A combination of general knowledge and deep specialization is what the state, as well as consumers, managers and citizens, expect of professionals. To gain full social recognition as an autonomous expert knowledge occupation, evaluation will have to make its way in the university world. Without further progress in this direction, evaluation will not secure the public franchise it needs to operate in the public sphere, secure the support of the state or elicit tolerance for its independence of decision makers.

Selectivity for Quality

Controlled access to the practice is a defining feature of professionalism. One source of opposition to systematic selection processes for access to the practice is rooted in the neoliberal notion that consumers should be free to hire anyone they wish, but evaluation is a public good. Poor-quality evaluation can destroy effective social programmes or give credence to misguided policy interventions.

Another rationale for resisting restrictions over entry is that it contributes to elitism and social stratification, but the alternative to thoughtful selectivity over membership in the evaluation profession has been tolerance of substandard-quality work. Ensuring that evaluators are equipped with the knowledge, skills and dispositions to exercise competent independent assessments of transformative social interventions is a collective responsibility that can only be satisfied through prudent access to the professional label.

The Self-Management Imperative

Finally, all contemporary models of professionalism stress the importance of self-management and autonomous control over occupational practices (Freidson 2001). Without professional autonomy, there is no collective accountability and no way to tap economies of scale in administration, manage the risks associated with weak standards or avoid capture of the occupation by vested interests or the state.

Self-management implies freedom in setting administrative rules, peer reviewing work quality, disciplining members and in extreme cases stripping them of their designation. These measures can be considered monopolistic, but the need to minimize the market disorder that inevitably prevails when quacks and amateurs can enter the fray with impunity and trigger distrust, confusion and poor service quality amply justify them.

Self-management would open up space to restore evaluation to the status of a vocation rather than a mere commercial enterprise; embed progressive, democratic values into its professional ethos; privilege inclusion and gender equality in its values framework; connect evaluation capacity-building initiatives with governance reforms and support professional development of young evaluators.

The Bottom Line

For evaluation to generate results and acquire influence, it will need to rise to the upper tier of the occupational ladder and acquire all the requirements

of professionalism. The four antecedents of professionalism sketched above are interrelated. Without demanding ethics, evaluation is easily captured and cannot be distinguished from other knowledge occupations. Without expert knowledge, evaluation would not add value to society. Without proven competencies and control over entry, evaluation quality is bound to be mixed. Without self-management, evaluation would remain submissive to vested interests. There is no shortcut.

Although they should do their utmost to manage the risks associated with the self-serving attitudes, elitist dispositions and monopolistic features of professionalism, evaluators will have to embark on an arduous professionalization journey to make a significant difference in the coming policy transition. Professionalization would facilitate widespread adoption of the necessary changes that the evaluation community has already endorsed. It would imply adoption of a demanding reform agenda focused on the demand side of evaluation practice. This is an exceptional transformation challenge, but these are exceptional times, and as Carol Weiss (1998, 325) famously opined, 'evaluation is not a stroll on the beach'.

References

- Bourguignon, François. 2016. Global versus national inequality. World Science Report. UNESCO, Paris, France.
- Bucy, Michael, Stephen Hall and Doug Yakola. 2016. Transformation with a capital T. *McKinsey Quarterly*. McKinsey and Company.
- Campbell, Donald T. 1971. Methods for the experimenting society, Paper presented to the Eastern Psychological Association, New York, and to the American Psychological Association, Washington, DC.
- Case, Anne, and Angus Deaton. 2020. The epidemic of despair. *Foreign Affairs* March/April. <https://tinyurl.com/5f96aesh>
- Chelimsky, Eleanor, and William R. Shadish, editors. 1997. *Evaluation for the 21st century: A handbook*. Thousand Oaks, CA: Sage Publications.
- Dahler-Larsen, Peter. 2012. *The evaluation society*. Stanford, CA: Stanford University Press.
- Datta, Lois-ellin. 2016. Who pays the piper: Funding sources and evaluation's contribution to equity. In: *Evaluation for an equitable society*, (eds.) S. Donaldson and R. Picciotto, 153–168. Charlotte, NC: Information Age Publishing.
- Davidson, E. Jane. 2005. Marketing evaluation as a profession and a discipline. Editorial. *Journal of Multidisciplinary Evaluation* <https://tinyurl.com/5h75ywwb>
- Davies, Ian C., and Julia Brummer. 2015. *Final report to the UNEG Working Group on Professionalization of Evaluation*. New York: UNEG.
- Economist. 2019. Democracy index 2019. <https://tinyurl.com/f262jusd>

- EvalPartners. n.d. EvalAgenda2020. <https://tinyurl.com/5cp8m97y>
- Freidson, Elliot. 2001. *Professionalism: The third logic*. Cambridge: Polity Press.
- Furubo, Jan-Eric, and Nicoletta Stame (eds.). 2019. *The evaluation enterprise*. New York: Routledge.
- Greene, Jennifer C. 2012. Values-engaged evaluation. In: *Evaluation for critical development results*, (ed.) Marco Segone, 192–206. New York: UNICEF.
- Guardian. 2019. The Guardian view on Davos: Elites without answers. <https://tinyurl.com/3tm8uz22>
- House, Ernest R. 2016. Inequality and conflict of interest in evaluation. In: *Evaluation for an equitable society*, (eds.) S. I. Donaldson and Robert Picciotto, 169–180. Charlotte, NC: Information Age Publishing.
- Kirkhart, Karen E. 2015. Unpacking the evaluator's toolbox: Observations on evaluation, privilege, equity and justice. *Evaluation Matters – He Take To Te Amoratawai* 1: 7–24.
- Klein, Naomi. 2007. *The shock doctrine: The rise of disaster capitalism*. Toronto: Knopf Canada.
- Mathison, S. 2016. Confronting capitalism: Evaluation that fosters social equity. In: *Evaluation for an equitable society*, (eds.) S. I. Donaldson and R. Picciotto, 83–107 Charlotte, NC: Information Age Publishing.
- Mertens, Donna M. 2005. Transformative paradigm. In: *Encyclopedia of Evaluation*, (ed.) Sandra Mathison, 422–423. Thousand Oaks, CA: Sage Publications.
- Milanovic, Branko. 2019. *Capitalism alone: The future of the system that rules the world*. Cambridge, MA: Harvard University Press.
- Mishra, Pankaj. 2018. *The age of anger: A history of the present*. London, UK: Random House, Penguin Books.
- Ofir, Zenda. 2019a. Competencies for post-normal evaluation, Part 3. Evaluation for Transformation blog. 14 September <https://tinyurl.com/y3a97ujd>
- Ofir, Zenda. 2019b. Towards Transformation, 3. A grand narrative for evaluation in this era. Evaluation for Transformation blog. 3 January. <https://tinyurl.com/dy496rbc>
- Parsons, Talcott. 1968. *Professions. International encyclopaedia of the social sciences*, vol. 12, 536–547. New York: The Free Press and Macmillan.
- Patton, Michael Quinn. 2008. *Utilization-focused evaluation*, 4th ed. Thousand Oaks, CA: Sage.
- Patton, Michael Quinn. 2010. *Developmental evaluation. Applying complexity concepts to enhance innovation and use*. New York: Guilford Press.
- Patton, Michael Quinn. 2020. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Picciotto, Robert. 2015. *Democratic evaluation for the 21st century evaluation*. Thousand Oaks, CA: Sage Publications.
- Picciotto, Robert. 2019. Donald T. Campbell's evolutionary perspective and its implications for evaluation. *Journal of Multidisciplinary Evaluation* 15 (33): 1–15.
- Pinker, Stephen. 2018. *Enlightenment now: The case for reason, science, humanism and progress*. London: Allen Lane.

- Revinga, Ana, and Meagan Dooley. 2019. Is inequality really on the rise? The Brookings Institution. <https://tinyurl.com/2ux9p8ux>
- Roser, Max. 2016. War and peace. <https://ourworldindata.org/war-and-peace>
- Roser, Max. 2019. Which countries achieved economic growth? And why does it matter? <https://tinyurl.com/2pehfeum>
- Roser, Max, and Mohamed Nagdy. 2013. Genocides. <https://ourworldindata.org/genocides>
- Roser, Max, and Esteban Ortiz-Ospina. 2013. Global extreme poverty. <https://ourworldindata.org/extreme-poverty>
- Roser, Max, Esteban Ortiz-Ospina and Hannah Ritchie. 2013. Life expectancy. <https://ourworldindata.org/life-expectancy>
- Roser, Max, and Hannah Ritchie. 2013. Homicides. <https://ourworldindata.org/homicides>
- Roser, Max, and Hannah Ritchie. 2019. Hunger and undernourishment. <https://ourworldindata.org/hunger-and-undernourishment>
- Schwandt, Thomas A. 2017. Professionalization, ethics, and fidelity to an evaluation ethos. *American Journal of Evaluation* 38 (4): 546–553.
- Schwandt, Thomas A. 2019. Post-normal evaluation? *Evaluation* 25 (3): 317–329.
- Scriven, M. 2007. The logic of evaluation. OSSA Conference Archive, University of Windsor. <https://tinyurl.com/4y775ark>
- Scriven, Michael. 2016. The last frontier of evaluation: Ethics. In: *Evaluation for an equitable society*, (eds.) S. I. Donaldson and R. Picciotto, 12. Charlotte NC: Information Age Publishing.
- Shadish, W. R., and J. K. Luellen. 2005. The history of evaluation. In: *Encyclopedia of evaluation*, (ed.) Sandra Mathison, 183–183. Thousand Oaks, CA: Sage Publications.
- Stake, R. E. 2016. Finding merit, diminishing hurtful discrimination, with less pursuit of meritocracy. In: *Evaluation for an equitable society*, (eds.) S. I. Donaldson and R. Picciotto, 181–198. Charlotte, NC: Information Age Publishing.
- Statistics Times. 2021. Projected GDP per capita ranking. <https://tinyurl.com/32mxht7f>
- Taleb, N. N. 2007. *The black swan: The impact of the highly improbable*. New York: Random House.
- United Nations. 1949. *World economic report 1948*. New York: United Nations.
- Vedung, Evert. 2010. Four waves of evaluation diffusion. *Evaluation* 16 (3): 263–277.
- Ventriss, Curtis. 2000. New public management: An examination of its influence on contemporary affairs and its impact on shaping the intellectual agenda of the field. *Administrative Theory & Praxis* 22 (3): 500–518.
- Wadhwa, Divyanshi. 2019. More men than women are literate. World Bank blogs. <https://tinyurl.com/2vxx9hxx>
- Weber, Max. 2002. *The Protestant ethic and the spirit of capitalism*. London: Penguin Books. Translated by Peter Baehr and Gordon C. Wells.
- Weiss, Carol H. 1998. *Evaluation*, 2nd ed. Upper Saddle River, NJ: Prentice Hall.
- Wike, Richard, Laura Silver and Alexandra Castillo. 2019. Many across the globe are dissatisfied with how democracy is working. Pew Research Centre. <https://tinyurl.com/dnjin9v8w>

PART IV

**THEMES
AND
CASES**





CHAPTER 12

Multiple Dimensions of Evaluation in Fragility, Conflict and Violence

INGA-LILL ARONSSON AND HUR HASSNAIN

Abstract. Evaluation in contexts affected by conflict and fragility is political and complex and can exacerbate violence. In such unpredictable environments, understanding how change happens is challenging because different actors at local, national and international levels have varied interests and definitions of what change is or may be. Reports on progress on the Millennium Development Goals (MDGs) show that these contexts lag in establishing robust monitoring and evaluation systems. Most examples of tools and resources available are from 'normal' contexts and do not fully support identification of the multiple biases at all levels during an evaluation exercise, including possible bias of the evaluator. In these contexts, collaboration with people on the ground is paramount to contextualize scenarios, tools and values in order to visualize what works or not for evaluation in a particular setting. The authors combine in this chapter background research completed, a reflection of their own personal experience, and the rich discussions with global evaluation practitioners who have worked in some of the worst conflict-affected contexts. During the IDEAS's Global Assembly in Prague, the authors conducted a one-day workshop on evaluation in fragile, conflict and violent contexts was attended by participants from many countries representing a range of leading actors and organizations that enriched this urgent topic.

Introduction

In this chapter, we critically reflect on some of the multiple dimensions of evaluations in fragility, conflict and violence (FCV) and in situations of global pandemic such as COVID-19, to *problematize* evaluation in environments that are fluid, complex, unpredictable or violent, with the aim of achieving transformational change within the broader frame of the Sustainable Development Goals (SDGs). Evaluation for transformational change is no easy task. As we argue here, it is even more difficult in FCV contexts, as well as during unstable times or crisis. In these circumstances, people are vulnerable; lack trust; live in complicated and sensitive relationships and lack faith in experts, outsiders, locals and government representatives alike. In a worst-case scenario, interaction with the wrong person could put one in danger or at minimum make it impossible to gather and process critical information and act on it, as the current pandemic has shown.

Hope, too, needs to be provided, and the Organisation for Economic Co-operation and Development (OECD) inspires it by commending the 2030 Agenda:

The Sustainable Development Goals set a roadmap for a better world. One where poverty, hunger, disease, climate change and gender inequality are no longer a threat to our planet and wellbeing. Instead, they chart a world where decent jobs for all, sustainable infrastructure, clean oceans and energy, responsible consumption and production, clean water and sanitation, and quality education, become the norm (OECD 2019, 3).

Most of us certainly want that better world that Agenda 2030 proposes. Nevertheless, from an evaluation perspective within FCV situations, the above quotation is problematic – because of its limitless optimism. This message would benefit from a touch of pessimism. From a realistic perspective, the SDGs might be difficult to achieve, partly because of the pandemic; Oxfam (2020) estimates that the crisis could push half a billion people back into poverty unless urgent action is taken to bail out developing countries. There is a risk that people may turn against the SDGs if they are perceived as unrealistic, or even belittling towards ordinary people, especially when considering the time frame of 2030. It may not be necessary to call the SDGs ‘worthless’, as Easterly (2015) did and as discussed in Van den Berg, Magro and Salinas Mulder (2019a), but there is a gap between the ‘real world’ and ‘real world evaluations’. This chapter aims at contributing with reflections on evaluation as a possible maker of transformational change, towards achieving the SDGs in the most difficult and confounding

circumstances, such as situations of FCV. The recently launched Global Evaluation Initiative, led by the World Bank and the United Nations Development Programme, estimates that only one-third of the countries who committed to achieving the SDGs have the monitoring and evaluation skills and capacities to do so (GEI 2020).

Evaluation in FCV is mostly – but not exclusively – associated with SDG 16, on peace, justice and strong institutions. Evaluation in FCV with respect to the SDGs must address complexity and unpredictability which are challenging in practice and theory in the evaluation field. This is not to claim that complexity and unpredictability are absent in normal circumstances, but in FCV contexts, specifically when transformations for the SDGs are a concern, this complexity and unpredictability are intensified, tighten and change quickly, and the stakes are high for all involved. Evaluators must get it reasonably right.

We began joint work on evaluation in FCV a few years ago, when we met at the 2017 IDEAS Global Assembly and realized that our different professional and geographical backgrounds offered us a fruitful platform for intellectual exchange and collaboration (see e.g. Aronsson and Hassnain 2019; forthcoming a; forthcoming b). This chapter will not offer a literature review on evaluation in FCV but rather will build on our experiences and the experiences of others with whom we had the opportunity to exchange ideas and perceptions in various evaluation fora using cases from the Global North and the Global South.

To lay the ground for critical reflections, we will begin by presenting a brief frame for FCV evaluation. We will then indicate challenges we faced evaluating in FCV and formulate a handful of guiding statements that summarize our first-hand experience. These were widely discussed and validated at the workshop that two of us facilitated at Evaluation for Transformative Change: Bringing Experiences of the Global South to the Global North, held in Prague in 2019¹. This conference was a joint effort of the IDEAS Global Assembly and the Third International Conference on Evaluating environment and Development.

Furthermore, by presenting a case of evaluation on violent extremism in Europe and a case on COVID-19 in Europe, we indicate that there are overlapping and distinct analytical levels between stable and unstable contexts that are worth exploring because we have not identified all the

¹ Multiple Dimensions of Evaluation in Fragility, Conflict, and Violence (EvalFCV). Workshop at IDEAS Global Assembly. Conference on Evaluation for Transformative Change. Prague, 1 October 2019.

mechanisms to be considered in FCV evaluations. Experts will probably not identify *all* factors involved, and some imperatives may vary from one situation to another. Nevertheless, a handful of precise indications are necessary and possible, as our experience in evaluation in FCV contexts and crisis show.

Collaboration is a key message in Van den Berg, Magro and Salinas Mulder's (2019b) book on transformational change. It is argued that collaboration is needed not only in the participatory methodology on the ground, but also between academic fields and evaluation units and between practitioners and theorists to make high-quality, viable progress towards the SDGs. At the same time, silos of knowledge on particular topics are essential building blocks of scientifically grounded work, which means that a required systems perspective and effective collaboration between experts from different areas of knowledge have not penetrated the evaluation field to a level that achievement of the SDGs require.

A Frame for Evaluation in FCV

It is estimated that 20 per cent of the world's poor are living in regions affected by FCV and that, by 2030, at least 46 per cent of the world's poor will be living in such regions (World Bank 2017) or more than 80 per cent if action is not taken (OECD 2018). These estimations are from before the COVID-19 crisis. Considering the many predictions being made, for example the Oxfam estimates discussed above, it seems likely that this pandemic will have dramatic effects on people living in countries affected by FCV and will push some additional states towards instability and fragility.

The SDG-16 Progress Report (Institute for Economics and Peace 2017) highlights that FCV-affected countries were on average 25 per cent more likely to have missed their Millennium Development Goals (MDGs) than other countries. It further explains that:

MDG indicators for which the majority of fragile and conflict-affected countries recorded the poorest results were those that addressed child mortality, maternal health and environmental sustainability. No conflict-affected country achieved the goal of reducing by two-thirds the under-five mortality rate between 1990 and 2015. Additionally, many of the fragile and conflict-affected countries have difficulty in maintaining the necessary systems to adequately capture the data. This can lead to poor quality data, resulting in situations appearing worse or better than what they are (Institute for Economics and Peace 2017, 7).

In evaluating the Aiding the Peace Initiative, Bennet et al. (2010) report that, in South Sudan, the support that multiple donors provided during 2005 to 2010 was often mistargeted. Reasons for this could be that donors did not fully take into account the key drivers of violence; essential services were overemphasized and security, policing and the rule of law were relatively neglected. With so many people living in vulnerable conditions worldwide, it is crucial that the right conditions be created to ensure rigorous, sensitive data collection and evaluation in FCV contexts, especially concerning achievement of the SDGs and to ensure informed decision-making in global pandemics (e.g. COVID-19).

Achievement of the SDGs in FCV states calls for a comprehensive repository of tools and resources to facilitate learning about different approaches to evaluating in FCV while making these evaluations climate and gender sensitive. Evaluation challenges in these contexts include, but are not limited to, difficulties in identifying and accessing affected populations because of rapidly changing context or moving populations; understanding power and relationship dynamics; facing fear and sensitivity regarding fact-finding missions and felt grievances; limited availability of good-quality data; paying attention to unintended effects; lack of appropriate tools and resources; and signs of corruption and human rights violations that are difficult to validate and report on. Other problems may include identifying competent evaluators who are willing to travel to conflict-affected areas and not being able to maintain impartiality throughout the evaluation process given the political context and difficulty in engaging with all key players (Hassnain et al. 2021). In addition, establishing indicators and their targets and measuring them are extremely challenging, for the many reasons discussed above. Peace, for example, is inherently political. There are multiple definitions of and perspectives on what peace is; it is perceived differently within and between countries and regions (Hassnain 2017). Furthermore, understanding and measuring the difference between targeted results and overall achievements in environments with FCV can be challenging. First, as discussed above, these states in FCV situations do not have sufficient resources for monitoring and evaluation, but even if they did, performance tracking systems that are mainly designed to measure the results of assistance in normal circumstances could easily miss progress – ‘Sometimes just keeping the lights on can be considered as a success’ in FCV environments (Kelly, Nogueira-Budny and Chelsky 2020). These issues are further described below.

Key Challenges of Evaluation

Programmes in contexts of FCV operate in a sensitive political environment and address complex challenges that are often hard to measure and report on. To evaluate them, the first task evaluators have is to build a comprehensive, systemic understanding of the context in which they are to work, including underlying cultural, social, economic and political factors and their interplay (Hassnain 2017; Aronsson and Hassnain 2019), but it does not end there. We have identified a variety of immediate challenges that evaluators must address:

- *Defining how change happens.* Evaluation in FCV contexts has weaker theoretical foundations and a limited evidence base of what works and what does not than in normal contexts. In such situations, programmes are founded on assumptions based on the principle that, if things work well, repeat, and otherwise drop them.
- *Fluidity.* FCV contexts are unpredictable, transitional in nature and fluid; hence it is difficult to submit them to a comparative analysis. Hassnain reports on his experience in Afghanistan in 2016, when many Afghans began to return from neighbouring countries. Many did not have a clear plan as to where to go to settle. Baseline data were gathered where they were initially located, but most moved to more convenient locations, such as where they had relatives or affiliations, before the intervention began. A similar situation occurred in South Sudan in 2017. These movements of beneficiary groups may mean shifts in operational priorities, resulting in limited or no reporting or sharing of such experiences while working on the ground. Aronsson reports on her experiences from the Zimapán resettlement project in Mexico, where counting affected people was difficult because they are constantly moving. Regardless of thoroughness, there always seemed to be an error at the end.
- *Risk of violence.* The imminence of bursts of violence limits face-to-face meetings for interviews and focus group discussions or even conducting surveys. Hassnain describes an experience in the Swat region of Pakistan in 2009, where a focus group discussion in a Taliban-affected community needed to be stopped because of the fear of escalating violence between the participants. Aronsson describes a violent uprising with gender dimensions in a resettlement project in central Mexico (Aronsson and Hassnain forthcoming a)

- *Trust.* Evaluations of FCV contexts and of peace, justice and strong institutions are fundamentally political and therefore fundamentally contested. Under different political perspectives, different things would be chosen to monitor and measure. Under those fluid, unpredictable circumstances, to reach agreement on indicators, tools and contextualized data is difficult and time consuming. In FCV contexts, everyone has their own agenda, and no one is neutral. To derive conclusions on the basis of sufficient triangulation efforts, as the best evaluation guides recommend, becomes of ultimate importance.
- *Learning.* What works today in a fluid context may not work tomorrow. In addition, there is limited or no culture of learning and sharing in such contexts because the contexts are changing rapidly and because of shifting priorities, low literacy levels, lack of trust between different parties and challenges related to data confidentiality. Documenting and reporting in such situations can be sensitive because of the politicization of international involvement and political sensitivities in national contexts, and evaluators may find it difficult to maintain a safe, credible evaluation space for learning and sharing.

The Guiding Statements

The following guiding statements provide a useful path for evaluators about to face a FCV situation. Although many of these statements may be valid in any evaluation, FCV contexts require that extreme attention be paid to each of them.

1. *The evaluand and the evaluator together determine the evaluation methods and approaches.* Do not get carried away by glittery tools and methods or smooth talkers.
2. *The context defines the methods.* Programmes in contexts of FCV operate in a political environment and address complex issues that are often hard to measure and report on. To evaluate them, the first task of evaluators is to have a comprehensive, holistic understanding of the context in which they work, including underlying cultural, social, economic and political factors and their interplay.
3. *People always come first.* The safety of respondents – individuals and institutions – evaluators, including enumerators, takes precedence over any kind of accountability measures. Consider seriously,

in advance, whether it is feasible and safe to conduct an evaluation in the given context.

4. *Be realistic and try to balance sensitive and objective knowledge and norms.* This is valid for both local and expert knowledge.
5. *Never reduce evaluation and social relationships in FCV contexts to a pedagogic or learning experience.* Follow an engaging, participatory process that makes a difference.
6. *The evaluators and enumerators must have documented ethical protocols, integrity and specific capacity and skills to work in FCV contexts.*
7. *Adaptive management approaches lead to iterative learning and context-responsive adjustments.*
8. *Participatory monitoring at regular intervals keeps the evaluators mostly informed of real-time changes on the ground, helping make sense of data over a given timescale.* Embedded evaluation or monitoring as a project activity, for which implementers are accountable and that is linked to feedback loops, may help the evaluator obtain data and make sense of it.
9. *Monitor for key factors identified through conflict analysis (e.g. dividers and connectors in conflict situations), but be prepared for participatory monitoring to elicit information on additional social impact issues, including gender and humanitarian protection.*
10. *Use information and communications technologies, mixed-methods and goal-free approaches* in FCV contexts and situations of pandemic in which the relationship between the causes and effects is complex (Hassnain and Lorenzoni 2020).

IDEAS's *Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners* (Hassnain, Kelly and Somma 2021) discusses these challenges in detail and asks important questions such as 'Who are the potential beneficiaries of the evaluation?' 'What are the unintended consequences/effects?' and 'What can be learned?' These questions may be addressed using mixed-methods, goal-free approaches such as *outcome harvesting*, working backward to the intervention logic and looking for *how* and *what* the intervention contributed. Moreover, using information and communications technology to prevent direct contact with people, where possible and feasible, is extremely helpful.

We agree with Chigas et al. (2006) that *trust* is a concept that could guide us in evaluation in fragility and conflict. Trust is universal, because without it, no social group could collaborate, which is part of the reason for

the success of humankind. Nevertheless, it is a morally and culturally constructed concept that could be misunderstood – or intentionally misused to manipulate people. In both cases, the evaluation turns into a pedagogic learning exercise more for the benefit of the evaluation team and the principals than for the intended benefactors. The evaluation commissioners and evaluators must find a fair balance by engaging as many local people as possible not only in collecting data in difficult situations, but also throughout the evaluation cycle, from study design to analysis and dissemination. This will not only make sure that the findings are relevant, but will also build local evaluation capacities.

Few professionals and academics would disagree with the above-presented dimensions of FCV evaluation. Its strength is the compilation and insights based on our experiences and on conversations with evaluators with or without experience in FCV contexts, but these identified dimensions of FCV must be further validated, especially concerning their use as instruments for transformational change.

A Pedagogy for FCV Evaluation Training

The participatory workshop we facilitated at the 2019 IDEAS Global Assembly gathered 25 evaluation experts from around the world, to whom we are indebted for the generous contribution through the exchange of lived experiences, ideas and knowledge. They represented bi- and multi-lateral agencies, including the United Nations, the World Bank, civil society organizations, think tanks and academia, from 23 countries, including some of the worst conflict-affected countries in the world, such as Afghanistan, the Occupied Territories of Palestine and Pakistan.

The workshop allowed us to take a step forward in our inquiry into evaluation in FCV states, which was a critical examination of the pedagogy for FCV evaluation. We explored insights, methods and theoretical implications for evaluation in FCV environments that can also be applied in situations such as global pandemics, where access to respondents is almost impossible.

Although one of our main objectives in facilitating this workshop was to introduce learning tools for enhancing practice for evaluation in FCV, relevant theoretical insights were also explored. The participants were eager to understand what, why and how things mattered in FCV in a different way than in other evaluation settings. Rather early in the workshop, participants and facilitators agreed that evaluation methods in FCV must be 'real

world evaluations' (Bamberger 2007) to adequately advise policymakers and practitioners on what works in the changing, complex contexts of FCV and, more importantly, how to avoid expensive mistakes. The challenges and guidelines presented in the previous sections were discussed and validated. Thereafter, a toolbox for evaluation in FCV was established in a spirit of joint understanding about participatory consultation, acceptance and agreement. This means that the tools are available, were shared with a relatively wide audience and can be used. Still, there is this nagging feeling that a roadmap must still be completed.

An implementation can fail for many reasons, but there is an interesting discontinuity between practitioners' practice on the ground and what they know and how the results are presented in the evaluation reports (especially the published versions). In the workshop, we found a wealth of tacit knowledge that workshop participants displayed in the dialogues but seemed to avoid articulating in the reports.

We also discussed the hesitation to raise, for example in focus group and individual interviews, sociopolitically sensitive topics that might be regarded as politically incorrect or were seen as a no-go zone for discussion. The idea of transformational change means that any cultural misguided sensitivity must be scrutinized and eliminated. Would this mean that the evaluator talks from a power position? To the contrary, the authors of this chapter reason that, to reach any transformational change, it is necessary to talk about difficult things. It is condescending not to discuss with the people involved, but such discussions demand considerable sociocultural competence and personal integrity. It would be a misjudgement to claim that all local evaluators possess the necessary social and cultural skills, as likewise it would be to claim that all external experts have competence, or for that matter, that the local affected people have it. Communication and collaboration are tools in a global interconnected world, but discussions must be honest and fearless; accountability, obligations and rights go hand in hand.

Based on the above, it can be argued that the FCV evaluation toolbox might lack aspects of relevant methods of FCV evaluations or that we are chasing something that can never be caught and presented in evaluation reports. There will always be a backstage and a frontstage. Here, theory enters the stage, because having a toolbox without adequate understanding of reality will only make us repeat ourselves in complicated hermeneutical circles or end up in power and identity politics cul-de-sacs or force reality into artificial boxes that makes no sense. As Gielen (2019) explains in her work on terrorism and political violence, theory-driven evaluation mainly examines the theory of change that contributes reasonable hypotheses but

fails to provide empirical testing. Process-oriented evaluation focuses on implementation of the intervention, whereas realist evaluation examines context-mechanism-outcome patterns (Gielen 2019, 1151). Gielen (2019, 1152) stresses that it is a question not of comparing methods but of avoiding falling into 'black boxes' and that non-linearity in these kinds of contexts should be realized and considered.

At the workshop, most of the learning took place in open-ended conversations on shared failures and successes based on a combination of people's experiences and newcomers' curiosity in a constant dialogue between participants and organizers. To claim that a modern learning environment must be open-ended and trustworthy is almost a platitude today, but it must nevertheless be said, because it can be confused with simplified social interaction learning that emphasizes that learning is all about an interactional space and relations between people and learning by doing, as taught for decades (Dewey 1938, Freire 1970). Moreover, according to Freire (1970) learning is mainly about the disruption of oppressive norms, which is reverberated in Bourdieu's 'symbolic violence' and cultural capital (Bourdieu 1977). Learning in this environment becomes all about power and identity, leading to identity politics and fragmentation of society. In this kind of learning environment, disagreement is rare, and discussions about uncomfortable topics are not welcome.

Instead, we addressed facts from practice-based cases and scenarios and connected them to relevant concepts in a factual learning environment as Christodoulou (2014) has promoted. We could have done more on that. We realized that there is much to explore when it comes to teaching and learning how to evaluate in the context of FCV. Although the current level of evaluation knowledge in the field of FCV is scattered and insufficiently tested, as mentioned above, constructivist learning methods, which are now re-evaluated, have dominated approaches to pedagogy for a long time, because the constructivist learning models stress social learning and power at the expense of factual learning. On the other hand, Freire's (1970) emphasis on oppression and deconstruction of norms is relevant for teaching about evaluation in FCV, because of the intricate sociopolitical and cultural-religious web of relationships between victims and perpetrators that often involve economic dependency. This web is 'context', which is a key concept but is so deep that it requires not only more time and trust than the present evaluation designs in FCV provide, but also another pedagogic approach if any transformational effects are to be achieved.

Learning methods for FCV evaluation must be rethought and tested. We feel that we must leave the 'nice learning environment', with its clear

rules, patterns, repetitions and immediate feedback, and move into a 'sneaky learning environment', with its unpredictable rules and open-endedness (Kahneman and Klein 2009; see also Jelmini 2020). This sounds harsh, but it would prepare evaluators for real-world encounters in which, for example, trust must be earned and critically examined if we aim at transformational change with strong participatory dimensions. International and local evaluators are triggers for transformation in these contexts, assuming that the affected people find it worthwhile and reasonable to participate and learn. We need to jointly develop a learning environment for evaluators in FCV. For example, we should neither give way to slack participatory methods such as a romanticizing of local knowledge nor use participatory methods to manipulate stakeholders. An honest, realistic learning paradigm is required if we are to leave no one behind and be inclusive, as Van den Berg, Magro and Salinas Mulder (2019a) argue. We must face any ideological conformity in our teaching methods; put the facts on the table and rely on a pedagogy that aims at rational inquiries, inspirational solutions and no safe spaces.

Furthermore, to improve learning methods, we must systematize evaluation in FCV and conduct a scientific review of presently used evaluation methods and theories, similar to Gielen's (2019) previously mentioned study about countering violent extremism. The authors of this chapter have initiated such a review in collaboration with the IDEAS thematic group on FCV, including involuntary resettlement.

Further Thoughts on Evaluation in Wicked Environments

The authors of this chapter claim that evaluation in FCV has a higher complexity level than evaluation under normal circumstances. This is not to be interpreted that we undervalue the complexity of evaluation in non-violent environments, but FCV is fluid, violent, dangerous and unstable, and evaluation in FCV takes place in politically charged environments.

Transformational change in development entails the following: 'peace requires a complete transformation from situations of conflict and violence to sustainable peaceful relations between warring factions and societies' (Van den Berg, Magro and Salinas Mulder 2019a, 6). This has been discussed in the peace and conflict research literature since at least the classic works of Galtung (1969) – a pioneer in the field. A signed peace agreement is just the beginning of a long process of reconciliatory measurements and projects before a possible sustainable peace might be achieved that can

be called transformational. Likewise, there are several 'realities' in a conflict that are interconnected in time and space, and the adversaries are heterogeneous (Kriesberg 2015, 7–11). Therefore, an understanding of *when* and under *what* circumstances a complete transformation takes place and *who* is involved. These elements are essential for FCV evaluation and evaluation research. The problem is that evaluation designs in FCV contexts, if tested in a given environment, may not always work in another.

Apart from the evaluation designs suggested previously, one possible evaluation design is Feinstein's suggestion of a dynamic evaluation with learning loops that he argues is necessary for transformation. The learning loop consists of four causally interrelated elements: dynamic evaluations of interventions → policy dialogue → policy change → transformational change → and back to the beginning. This loop assumes that transformational change will be triggered and accountability supported (Feinstein 2019, 26). First, it is not clear where in the societal structures this transformational change occurs. Second, Feinstein assumes that policy matters and that a trickle-down effect occurs during implementation. This is a problem. In involuntary resettlement, policy safeguards (e.g. the World Bank's) have been applied since the 1980s, but research evaluation has shown that there is a worrisome gap between policy and implementation. The reasons are multidimensional.

The authors of this chapter are concerned that, in this cloudy nexus of relationships in FCV, there is a real risk of harm because of the fluidity and blurriness of context and stakeholders. Coherency, in the sense of systematic connection between parts, is a main difficulty in FCV evaluations. Maybe the only regularities are ad hoc events, fluid contexts and solutions, and if that is the case, evaluation must learn how to address those kinds of processes in a systematic way. This is not only a methodological and theoretical question, but also a question of legitimacy. It is about the use of resources for implementation of the SDGs in a meaningful way, as formulated by the OECD (2018, 13–14).

The COVID-19 Global Pandemic and FCV Evaluation

Hassnain et al. (forthcoming) argue that the COVID-19 pandemic is a case of FCV evaluation because the economic downturn of lower- and middle-income countries is likely to lead to an increase in violence and political instability. The Institute for Economics and Peace, in its report on

COVID-19 and Peace (2020), stated that most indicators in the Global Peace Index are expected to deteriorate. The one area that may improve is military expenditures, as countries redirect resources to propping up their economies.

The theme of the IDEAS Prague Conference was sharing experiences between the Global South and the Global North, which is in line with the idea of collaboration in FCV evaluation to enhance capacity, and this is why we bring cases from the Global North and the Global South.

In the Global South, the coronavirus has in many instances united people in raising their voices and coming out onto the streets in protest of how the states are handling such cases. For examples, millions of Brazilian protesters in São Paulo and Rio de Janeiro demonstrated against the government's handling of the coronavirus pandemic. They protested and called for the president to step down.

Prison breaks were reported in Venezuela, Brazil and Italy, with inmates reacting violently to new restrictions associated with COVID-19. Drug trafficking and other types of crime have seen a temporary reduction as a result of social isolation around the world, but reports of domestic violence, suicide and mental illness have increased (Institute for Economics and Peace 2020).

In the Global North, the scenario has been different. For example, in Sweden, the authorities did not lock down the country but decided to rely on *individual responsibility*. People were asked to work at home, avoid public transportation, socially distance, wash hands and stay at home if they felt sick. No punishments or reprisals were given. The strategy was intended to flatten the curve of infection so that the medical system would have time to adapt. It might also have encompassed the idea of achieving herd immunity for the population. This strategy might have resulted in a higher percentage of deaths than in neighbouring Nordic countries². Future evaluation research will show whether this was the case. As in the rest of the world, the distribution of people with COVID-19 is uneven, and state epidemiologist Anders Tegnell stated that 'This is an illness with very strong socio-economic links' (Omni 2020). The detailed reasons for this will be analysed in future

² Folkhälsomyndigheten (Public Health Agency of Sweden) provides the country with daily statistical updates and information on the distribution of documented COVID-19 deaths and the number of infected people distributed over the country according to region, urban versus rural districts, and male versus female. The statistics are public and transparent. Calculation mistakes are acknowledged and corrected in public, as is the fact that COVID-19 test kits imported from China have been inadequate. Affected people are informed and asked to test again (Folkhälsomyndigheten 2020).

research. Here we will only conclude that Sweden is not an FCV evaluation context because it has functioning multi-actor policy processes in place that intercepted the problems, one after another, and a large, educated population that more or less followed the rules. These multi-actor policy frames have been built up over a long time and are resilient.

On the other hand, an FCV context lacks stability, although stability per se also needs further analysis, as discussed in the peace and conflict literature. The lack of stability (instability?) is combined with a small, educated elite and a large uneducated population who are used to corruption, lack of transparency and government incompetence, or at least a perception that this is the case. Without stability (sustainable socioeconomic-political and cultural structures) and trust, a country faces immense difficulties in the fight against a pandemic.

An effective FCV evaluation must explicitly address these issues in an unsentimental, non-ideological, rational way and not uncritically lean on well-meant general concepts repeated over and over in research and evaluation. For example, let us bring up the concept of 'local', which is usually approached as a level but maybe is better seen as 'a standpoint based in a particular locality, but not bounded by it', as by Shaw and Waldorf (2010, 6) suggested. This is context, and we must cut through its complexity. Maybe a way forward is, as Feinstein (2019, 21–25) argues, a dynamic evaluation with its key aspects of relevancy, multiple methods, scaling up, quasi-real-time evaluation and political sensitivity. The dynamic approach is valid because of the fluidity of an FCV context, and further experimentation to assess its compatibility with the suggested methodological elements is needed. Alternatively, what might be needed is a system thinking evaluation approach (Magro and Van den Berg 2019) that can incorporate and make use of fine-meshed qualitative data in a grid of quantitative data. There is no use in continuing with 'meaning-making' studies in an FCV context, if they continue to fail to operationalize their important messages, especially in an FCV context. It could even be dangerous. In the following section, we will discuss qualitative data that are systematized with the help of the realistic evaluation model using key concepts.

Violent Extremism and Evaluation Methods

Exploring further the authors' claim that evaluation in FCV is special because of the context, we will examine closely a case of evaluating programmes for countering violent extremism in Europe. This case study could help identify

possible further criteria for FCV evaluations or, should we say, evaluation in FCV contexts.

Gielen (2020), in her doctoral thesis 'Cutting Through Complexity', has investigated how to evaluate countering violent extremism in Europe. In two earlier articles (Gielen 2018; 2019), she developed and confirmed her argument. Here we will focus on the article on how to evaluate female jihadist exit programmes in the Netherlands (Gielen 2018).

When the Islamic State in Syria (ISIS) proclaimed its caliphate in June 2014, it led to an increase in female recruitment from European countries – young women who travelled to Syria and Iraq to marry ISIS fighters. They became known as 'jihadi brides' to become mothers to 'cubs of the caliphate'. In Europe, they were seen as victims, and Gielen (2018) argues that this 'victimization' prevented effective policies against female violent extremism. Recent empirical studies have shown that female jihadists are far from victims, and this requires a more nuanced view of the returning female jihadists in order to develop effective counter-measure programmes of deradicalization and reprogramming. These programmes must be tailor-made to be successful, and this is a challenge to prevailing evaluation methodologies. Gielen suggests that realistic evaluation is an option to find out what works for whom and when. This is also in line with our experiences with FCV, as discussed earlier.

Exit programmes consists of several elements, such as deradicalization (changing extremist beliefs), disengagement (dissuading from violent extremist action), reintegration and rehabilitation. All of these elements, if implemented, can lead to transformational change for the individual and for society.

Their deconstruction, which starts with their binary opposition, makes the elements mentioned above operational; for example, deradicalization starts with an investigation of radicalization, in which push and pull factors were identified, such as roles and gender (Gielen 2018). Female jihadists' gender roles were often underestimated, and they became an unrecognized and dangerous threat. Any exit programme must consider this.

The problem with the existing exit programmes was that, as Demant and colleagues say, 'exit programmes for jihadists focus too much on normative factors, concentrating on theological and ideological issues, and as a consequence overlook affective factors such as the family and peer network' (quoted in Gielen 2018, 460). This is also in agreement with our approach to FCV – to leave the frontstage context and reach the backstage context to achieve a kind of holistic view while still keeping the ethnographic details in mind.

Furthermore, Gielen (2018) argues that the local level was crucial for the understanding and success of the exit programme. Here, 'local level' means municipalities' multi-agency resources and management, such as legal, administrative and 'soft' resources (psychological counselling, family support, practical support with job and housing) and help with breaking with extremist networks, because grooming is one of the biggest problems. When a female jihadist returns, she is contacted within a short time (48 hours) to be convinced or threatened to rejoin the network.

Hence, 'local' comes back into the evaluation model as a key concept and engages with multi-actor agencies in a processual way. In Gielen's final model of a successive exit programme for female jihadists, several well-established, key elements return. She stresses a dense contextual approach. The municipality and its institutional setting constitute the frame of this context. Within this frame, gender perspectives and demography are taken into account and linked to mentoring programmes to promote trust and establish a long-term, stable environment. The stability depends on the quality of the relations between the agencies in charge of the case. Furthermore, concrete legal measures are applied to prevent the women from rejoining the network. Such legal measures can include confiscating passports, prohibiting contacts and banning social media (Gielen 2018). Pawson and Tilley (1997, 40) argue that, to establish an effective exit programme, multi-method data collection is necessary but without falling into 'the experimentalist trap to compare "inputs" and "outputs" in the sense that some programmes do work and others do not. Rather, realist evaluation should concern itself with the "make-up" of the interventions and respondents to address the question why some programmes work better for some than for others.'

Whether the above reasoning would effect transformational change is an open question, but it might be, as Feinstein (2019) argues, that to be transformative, evaluation must change focus from projects and programmes to strategies and policies. In addition, gender researchers argue that systemic gender knowledge is a presumption of transformational change, because 'scholars agree that gender inequality is systemic and that participants in gender equality interventions need knowledge on gender inequality processes' (Lansu, Bleijenberg and Benschop 2019, 1589). From an FCV evaluation perspective, we do not know what the triggers are for transformational change in any systematic way, and we are painfully aware of the gap between policy and implementation. Even more difficult, we have not pinpointed an exact moment in time when we can say whether an intervention has caused more harm than good.

Conclusion

Evaluation, like development aid, can unintentionally exacerbate tensions in ways that harm conflict-affected populations if care is not taken to develop and integrate layers of sensitivity into the design and approaches and throughout the evaluation cycle.

The contexts of FCV pose particular challenges for evaluation, such as difficulties in accessing the affected population; limited availability of good-quality data; lack of appropriate tools and resources; challenges in navigating conflicted stakeholder and informant relationships; high levels of unintended effects; signs of corruption and human rights violations that are difficult to validate and report on and political volatility that makes it difficult to identify key stakeholders of the intervention to be evaluated. These difficulties, if not properly addressed or mitigated, can call the validity of the evaluation into question. Mitigating them can be an important challenge and sometimes requires redefining the direction, purpose or scope of the evaluation, which requires that not only the evaluation team, but also the donor or funding agency be flexible.

Although evaluation data collection methods continue to evolve, there is a need for a more comprehensive repository that assembles guidance from different sources to facilitate learning about different approaches to conducting evaluation in the context of FCV in climate- and gender-sensitive ways. These are politically volatile environments, and it can be difficult to identify the key stakeholders that must be on board for an authentic evaluation process because the narration and understanding of conflict is different for each party or institution. It can be challenging to find the appropriate direction, purpose and scope for the evaluation, and sometimes this has to be redefined when in the field.

In FCV contexts, the importance of understanding the cultural, socio-economic and political context assumes a higher importance given the inherent complexities of such contexts. This includes not only facts about the conflict, but also the culture, economy, sociopolitical structures and safety – understanding in its true sense. This may be impossible with the present routines of evaluation missions. Nevertheless, this is what we have to work with. The authors emphasize the importance of conducting an in-depth desk study, including a detailed evaluability assessment to consider whether the evaluation is feasible and appropriate in the given context. This is also helpful in many ways to clarify data gaps and any other operational evaluation matters. In addition, the MDG progress reports indicate that data collection in FCV environments has been poor or that data

on the indicators are unavailable on numerous occasions. Some of the key reasons behind this is the absence of evaluation systems, tools and opportunities. The Global Evaluation Initiative verifies this need to harmonize tools and resources for effective evidence generation at all levels, indicating that only one-third of the countries that committed to achieving the SDGs have sufficient monitoring and evaluation capacities to measure any changes on the ground.

The bottom line is that evaluating in FCV contexts always starts with investing time and resources in analysing the context, including the key drivers of conflict and how these affect the rest of the society, including the interventions designed in such contexts, and vice versa. It will not prevent evaluations from missing or poor-quality data or, more importantly, from making expensive mistakes. We must learn to be imperfect, although we must strive towards perfectness and, foremost, help evaluators ensure learning and promote accountability at all levels. Just beware of the fact that, in normal circumstances, failures bring learning, whereas failure in proper planning and executing an evaluation may mean serious life-threatening consequences.

Let us conclude with some final thoughts on these intricate issues. The impression is that evaluation for transformative change envisions an Eden world of truth, equality, justice and peace, echoing Bob Thiele's famous 1967 song, magnificently interpreted by Louis Armstrong, 'What a Wonderful World'. The problem is how to reach this wonderful world, and the path is stacked with obstacles. As we see it, transformation is about memory but also about forgetting, and maybe (but only maybe) this is particularly important in fragile and violent societies. Transformation is about how and what to remember and how and what to forget. To reconcile, or at least to go on, is to be able to hear the sound of things falling, in the words of Juan Gabriel Vásquez (2011). It is also about not repeating 'This way for the gas, ladies and gentlemen', following the book title of holocaust survivor Tadeusz Borowski (1967).

Feinstein (2019) is referring to transformation in his review of the famous novel *Il Gattopardo* (*The Leopard*), by Giuseppe Tomasi di Lampedusa, published posthumously in 1960. Feinstein takes Lampedusa's famous quotation – 'Everything must change for everything to remain the same' – literally (to change to be the same) and regards it as a micro-event that does not really influence the larger societal change. We think this is a misinterpretation, because on Don Fabrizio's deathbed, the quotation becomes 'if you don't change, time will change you' (di Lampedusa 1960), or in the words of the Museum of di Lampedusa, 'It is a novel where the temporal

limits of human nature are always present, melancholic, touching and wise' (Butera 28 Apartments 2020). As we see it, transformation is about temporal and spatial orders and the human consciousness.

Evaluating for transformational change in unstable and fragile circumstances is to cut through complexity. We believe that evaluation and evaluation research must be a player in this matrix, but we need to recognize, grasp and learn how to use fewer tangible elements of change to make it happen, as the authors of this chapter have discussed in an earlier text (Aronsson and Hassnain 2019) on value-based evaluation for transformative change.

References

- Aronsson, I-L., and H. Hassnain. 2019. Value-based evaluations for transformative change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R.D. van den Berg, C. Magro and S. Salinas Mulder, 89–104. Exeter, UK: IDEAS.
- Aronsson, I-L., and H. Hassnain. Forthcoming a. Ethics in evaluation in resettlement related interventions. In: *Evaluation in an ethically challenged world*, (eds.) R.D. van den Berg, P. Hawkins and N. Stame. Exeter, UK: IDEAS.
- Aronsson, I-L., and H. Hassnain. Forthcoming b. Evaluation and ethics in contexts of fragility, conflict and violence. *Evaluation in an ethically challenged world*, (eds.) R.D. van den Berg, P. Hawkins and N. Stame. IDEAS
- Bamberger, M. 2007. *RealWorld Evaluation: Working under budget, time, data and political constraints*. Oregon Program Evaluators Network Brown Bag Series, 27 September.
- Bennett, J., S. Pantuliano, W. Fenton et al. 2010. *Aiding the peace: A multi-donor evaluation of support to conflict prevention and peacebuilding activities in southern Sudan 2005–2010*. Hove, UK: Itad Ltd.
- Borowski, Tadeusz. 1967. *This way for the gas, ladies and gentlemen*. London: Penguin Books.
- Bourdieu, Pierre. 1977 [1972]. *Outline of a Theory of Practice*. Trans. Richard Nice. Cambridge, UK: Cambridge University Press.
- Butera 28 Apartments. 2020. The Leopard – Giuseppe Tomasi di Lampedusa. Webpage. <https://tinyurl.com/4sbw3exu>
- Chigas, Diana et al. 2006. *Has peacebuilding made a difference in Kosovo? A study of the effectiveness of peacebuilding in preventing violence: Lessons learned from the March 2004 riots in Kosovo*. Case Study, Cambridge, MA: CDA Collaborative Learning Projects and Care International. <https://tinyurl.com/4yawunck>
- Christodoulou, D. 2014. *Seven myths about education*. London, UK: Routledge.

- Dewey, J. 1938. *Experience and education*. West Lafayette, IN: Kappa Delta Pi.
- Di Lampedusa, G.T. 1960. *Leoparden (Il Gattopardo)*. Swedish translation. Bonnier: Stockholm.
- Easterly, W. 2015. The SDGs should stand for senseless, dreamy, garbled. *Foreign Policy* 28 September. <https://tinyurl.com/y21mtp2x>
- Feinstein, O. 2019. Dynamic evaluation for transformational change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R. D. van den Berg, C. Magro and S. Salinas Mulder. Exeter, UK: IDEAS.
- Folkhälsomyndigheten (Public Health Agency of Sweden). 2020. COVID-19 statistics. <https://tinyurl.com/49d5a4jr>
- Freire, P. 1970. *Pedagogy of the oppressed*. Penguin Classic.
- Galtung, J. 1969. Conflict as a way of life. *Essays in Peace Research* 3: 484–507.
- Gielen, A.-J. 2018. Exit programmes for female jihadists: A proposal for conducting realistic evaluation of the Dutch approach. *International Sociology* 33 (4): 454–472.
- Gielen, A.-J. 2019. Countering violent extremism: A realist review for assessing what works, for whom, in what circumstances, and how? *Terrorism and Political Violence* 31 (6): 1149–1167.
- Gielen, A.-J. 2020. Cutting through complexity. Evaluating countering violent extremism (CVE). Doctoral thesis. University of Amsterdam.
- GEI (Global Evaluation Initiative). 2020. Global Evaluation Initiative. <https://www.globalevaluationinitiative.org/>
- Hassnain, H. 2017. Measuring peace in our lifetime: Evaluating SDGs. *Alliance Magazine*. <https://tinyurl.com/yfmbh5>
- Hassnain, H., L. Kelly and S. Somma, eds. 2021. *Evaluation in contexts of fragility, conflict and violence: Guidance from global evaluation practitioners*. Exeter, UK: IDEAS.
- Hassnain, H., and M. Lorenzoni. 2020. Embracing the pandemic. Institute for Economics and Peace. 2017. SDG 16 progress report. Sydney.
- Institute for Economics and Peace. 2020. COVID-19 and Peace. Sydney.
- Jelmini, M. 2020. Tidig specialisering inte recept på framgång – snarare tvärtom. *Svenska Dagbladet* 13 August 2020, 26–27.
- Kahneman, D., and G. Klein. 2009. Conditions for intuitive expertise. A failure to disagree. *American Psychologist* 64 (6): 515–526. <https://doi.org/10.1037/a0016755>
- Kelly, L., D. Nogueira-Budny and J. Chelsky. 2020. Conflicting results: Measuring outcomes in situations of conflict. World Bank blogs. <https://tinyurl.com/s2fennk>
- Kriesberg, L. 2015. *Realizing peace. A constructive conflict approach*. Oxford: Oxford University Press.
- Lansu, M., I. Bleijenberg and Y. Benschop 2019. Seeing the system: Systemic gender knowledge to support transformational change towards gender equality in science. *Gender Work Organization* 26: 1589–1605.

- Magro, C., and R.D. van den Berg. 2019. Systems evaluations for transformational change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R.D. van den Berg, C. Magro and S. Salinas Mulder, 131–152. Exeter, UK: IDEAS.
- OECD (Organisation for Economic Co-operation and Development). 2018. *States of fragility 2018*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264302075-en>
- Omni. 2020. Undersökning: Snabbast spridning i utsatta områden. <https://tinyurl.com/3z648np9>
- Oxfam. 2020. Oxfam report. <https://tinyurl.com/56vmc8th>
- Pawson R., and N. Tilley. 1997. *Realistic evaluation*. London: Sage. <https://tinyurl.com/rcazb76v>
- Shaw, R., and L. Waldorf, with P. Hazan (eds.) 2010. *Localizing transitional justice. Interventions and priorities after mass violence*. Stanford, CA: Stanford University Press.
- Van den Berg, R. D., C. Magro and S. Salinas Mulder. 2019a. Introduction: Bringing challenges for evaluation for transformational change from Guanajuato to Prague. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) R.D. van den Berg, C. Magro and S. Salinas Mulder. Exeter, UK: IDEAS.
- Van den Berg, R. D., C. Magro and S. Salinas Mulder (eds.). 2019b. *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*. Exeter, UK: IDEAS.
- Vásquez, J.G. 2012. *The sound of things falling*. New York: Riverhead Books.
- World Bank. 2017. Fragility, conflict and violence. <https://tinyurl.com/3tf7a3hd>



CHAPTER 13

The Potential of Value Chain Development for Transformative Change: The Experiences of AfDB and IFAD

FABRIZIO FELLONI AND GIRMA KUMBI

Abstract. This chapter examines the contribution of the African Development Bank and International Fund for Agricultural Development to agriculture-related value chain development, based on evaluations that these organizations conducted. The chapter offers a systemic perspective from which to conceptualize value chains and value chain development for poverty reduction. If well designed and implemented, value chain support can lead to transformative changes for smallholder farmers and rural small-scale producers, but both evaluations conclude that working on value chains requires major changes in the organizational culture. This chapter emphasizes the importance of corporate-level strategies in creating consistency and guidance on value chains and thereby assisting with project design and implementation. Evaluation findings indicate that reaching impoverished rural farm households through value chain approaches requires specific attention. Having approached the topic of value chains from a system perspective, this chapter identifies five key fundamentals and enablers that characterize successful agricultural value chain development, highlights policy implications and makes key recommendations. It provides some lessons that will be relevant to future evaluations on this topic.

Introduction

Background

Large-scale processing, wholesale and logistics operations serving retailers, foodservice operators and large markets have increasingly been replacing traditional food systems through value chains. Small-scale producers are still responsible for a large part of food production in the world but receive a disproportionately low share of its market value¹. Governments, development agencies, non-governmental organizations and some private companies have begun showing interest in making food value chains more socially inclusive and environmentally responsible. In addition, the 2030 Agenda for Sustainable Development has focused on the principle of 'no one left behind'. This aphorism brought attention to the topic of inclusiveness, the ability of poor producers and other marginalized groups to participate in value chains without increasing inequality. The expectation was that supporting value chain development in an inclusive manner would bring about a transformative change for small producers by enabling better contractual conditions and ultimately access to a larger share of the final consumer price.

This chapter provides an overview of the findings of two recent evaluations that the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD) and the Independent Development Evaluation of the African Development Bank (AfDB) conducted (IFAD IOE 2019 and AfDB IDEV 2018, respectively). Both are international financial institutions providing financing to governments and non-sovereign entities for preparation and implementation of development projects. IFAD specializes in rural development and poverty alleviation. AfDB's portfolio spans several sectors, but approximately 11 per cent was dedicated to agricultural development in 2016.

Marked Growth in the Financing Portfolio Relevant to Value Chain Development

At IFAD, interest in and commitment to developing or improving pro-poor value chains have grown significantly since the mid-2000s. This was intended to mark a departure from the previous almost exclusive focus on production. It started from the issuance of its Strategic Framework

¹ In 2013, it was estimated that smallholder farmers produced up to 80 per cent of food in Asia and sub-Saharan Africa (Arias et al. 2013).

for 2007–2010 and continued through the Strategic Frameworks for 2011–2015 and 2016–2025. The proportion of value chain-relevant projects approved increased from 41.5 per cent between 2007 and 2009 to 72.3 per cent between 2016 and 2018, and the proportion of loan volumes with value chain components increased from 50 per cent to 81 per cent over the same period.

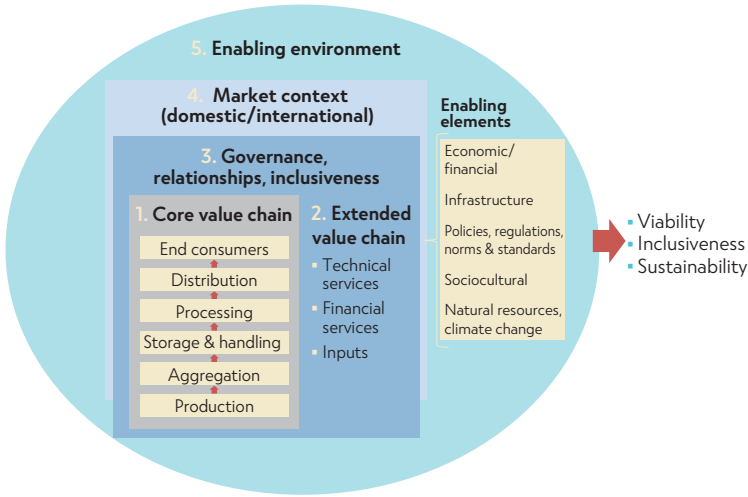
AfDB's Ten-Year Strategy (2013–2022) expresses the Bank's ambition to assume a more central role in Africa's development. With respect to agriculture, this strategy places a more direct focus on achieving food security through increased production or access to disposable income for purchase of food. The Feed Africa Strategy (2016–2025) aims to transform African agriculture into a competitive, inclusive agribusiness sector that creates wealth, improves lives and secures the environment. The Feed Africa Strategy promotes an integrated value chain development approach, with the private sector at the heart of the development process. It also envisages that the public sector will facilitate investments in the agricultural sector, particularly when serving smallholders and small and medium-sized enterprises. Inclusiveness is important to ensure that benefits from value chain development reach poor farmers, women and young people. AfDB's project interventions focusing on value chain development have increased from 15 per cent during 2005 to 2010 to 52 per cent during 2011 to 2016.

A Systemic Representation of a Value Chain

In the literature, the usual definition of value chain is the set of units of production and processing along the chain of activities required to bring a product from the initial input supply stage through the various phases of production and processing to its final market destination (e.g. Kaplinsky and Morris 2002)². This definition does not take into account the complexity of a value chain, its embeddedness in a market system, the importance of an enabling policy environment and the conditions for a value chain to develop in an inclusive manner. It is more useful to adopt a systems approach and consider a value chain as a system, of which the supply chain is only a subsystem that is connected to other subsystems (figure 13.1, subsystem 1). The supply chain subsystem comprises a series of functions from production to aggregation, storage and handling, processing, and distribution, finally reaching end-users (FAO 2014; M4P and DFID 2008; USAID 2014). An

² The term 'value chain' is credited to Michael Porter (1985).

Figure 13.1 Representation of a Value Chain System



Source: IFAD IOE (2019), adapted from FAO (2014), with inputs from M4P (2014) and USAID (2014).

additional subsystem (subsystem 2), which is ignored or downplayed in many schematic representations, comprises providers of goods and services such as inputs to production (e.g. seeds, fertilizers), financial services, advisory services and market information.

Part of a value chain system is its governance (subsystem 3), which refers to how business linkages are structured along the chain and to the relationships among the stakeholders, including buyers, sellers, service providers and regulatory institutions. For value chains that cut across national borders, governance may be particularly complex because stakeholders are located in different countries and subject to different policies and regulatory provisions.

Governance is essential for inclusion of the poor, given that one of their most frequent problems is lack of power and voice in the system. Strengthening their representation and bargaining power can increase the economic and non-economic benefits they receive, such as through building the capacity of small producers to negotiate terms of trade with buyers.

A value chain also interacts with a market (subsystem 4), which is characterized by the interaction of supply and demand (local, national or international), a set of regulations and the level of competition between

stakeholders (or varying degree of monopolistic power). The enabling environment (subsystem 5) determines to what extent a value chain favours the flow of commodities, money and information in a viable manner in the short term; is sustainable in the long run and generates equitable outcomes for its stakeholders.

The systemic representation of the value chain is a useful conceptual reference for those in charge of designing programmes and those that evaluate them. The most important lesson learned is the interconnectedness between the subsystems within the broader value chain system. Too often, the value chain is identified in a narrow manner – with the supply chain and the importance of governments, markets and regulations. Project designs do not need to cover all the subsystems, and in many cases, it may be too ambitious to do so, but they need to be cognizant of the system complexities, even if they are only intervening in a single subsystem or parts of it, at least as guidance to prioritize their planned activities. As discussed further below, the initial drive towards value chain development at IFAD and AfDB was not based on a systemic value chain perspective.

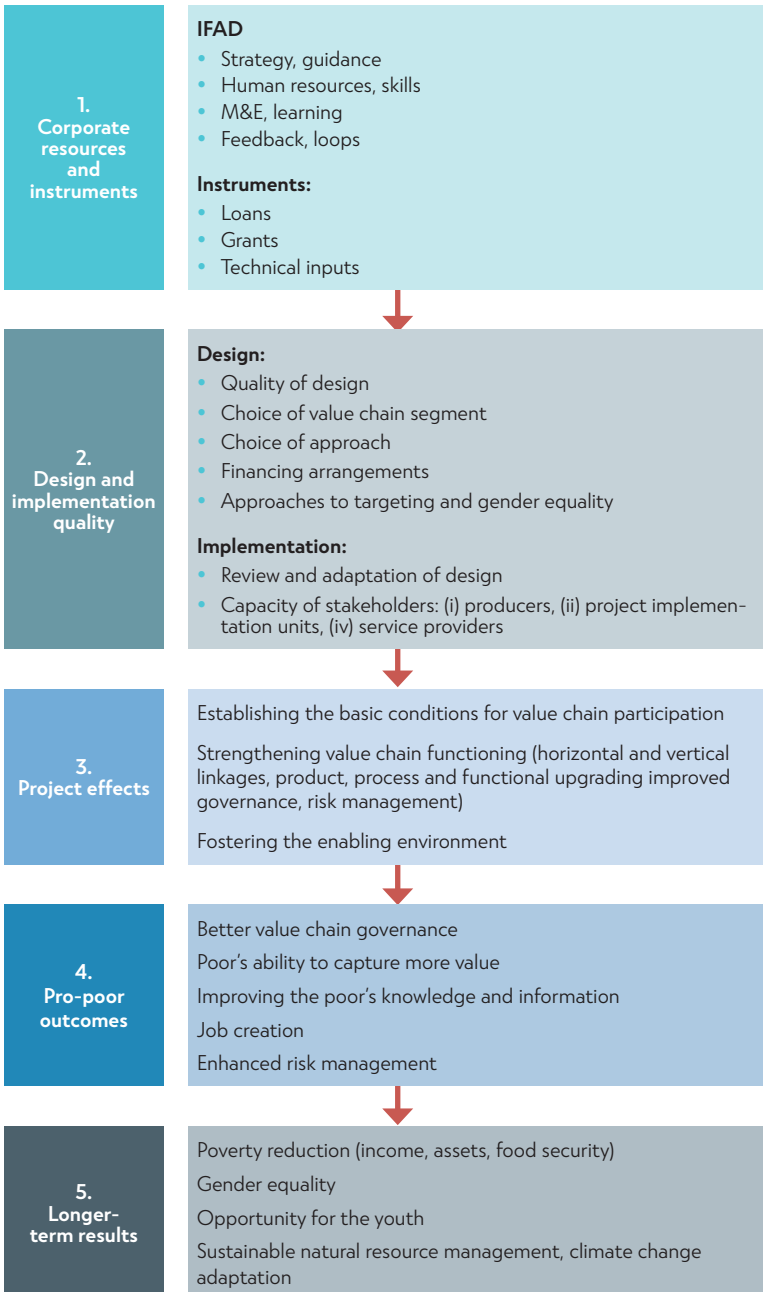
Highlights of the Methodology Used in the AfDB and IFAD Evaluations

Although the two evaluations addressed different corporate mandates, institutional contexts and business models, there were similar fundamental questions, such as:

- Were the organizational setup and instruments conducive to supporting value chain development?
- Have the strategies and interventions been relevant in their focus on value chain development?
- To what extent have value chain development interventions been effective in achieving their planned objectives and the corporate mandate?
- Have value chain development interventions been inclusive (e.g. of the poor, women and youth)?

Figure 13.2 illustrates that the evaluation first explored IFAD's organizational capacity to promote pro-poor value chains. It reviewed corporate resources and instruments to support governments and other country partners in value chain development, quality of project design, implementation

Figure 13.2 IFAD’s Support to Value Chain Development for Poverty Reduction



Source: Adapted from IFAD IOE (2019).

performance, immediate project effects and project outcomes for the poor and longer-term results.

Main Findings

Corporate Organizational Aspects

Supporting agricultural value chains was expected to have transformative effects for rural small-scale producers. Experience showed that the development organizations financing these programmes needed to transform how they operate.

According to IFAD (IFAD IOE 2019), although the number and financial volume of investments with value chain elements increased significantly from 2007 to 2017, the organization did not elaborate a strategy, a policy or any comprehensive corporate guidance on value chain development. Such strategy or guidance, building for example on a systemic approach to value chains, could have built consensus on how value chain support relates to the mandate of poverty reduction and, in particular, through what channels and mechanisms poor people could benefit. There was, initially, limited emphasis on training of staff on the concepts of value chains and value chain development and on making explicit the nexus between being pro-poor and inclusive. Value chain development also implies collaborating with private entrepreneurs and companies, which was a relatively new concept at IFAD.

Government agencies execute IFAD-funded projects, which are staffed with employees that, largely, come from the public sector and have experience in agricultural production, civil engineering, procurement and project administration. Value chain development requires new skills and a business perspective. The need for value chain or marketing specialists was only occasionally anticipated in project management units. Numerous project managers had limited familiarity with value chain development. There was no capacity-building strategy through which technical support opportunities were defined in a coordinated manner and synchronized with project activities. On a positive note, IFAD staff and project managers displayed willingness to adapt, experiment and learn, although interpretations varied widely as to what supporting value chains meant and how rural poor people should be engaged.

A longitudinal review across generations of project designs showed that there was an evolution in project conceptualization. Whereas IFAD-funded projects formulated until the first half of the 2000s were typically focused on improving primary production, with time, the marketing of products and

concerns about 'good prices' and selling opportunities for small farmers and producers had come to the forefront of project formulation. The evaluation also found considerable 'learning by doing'. Projects with better value chain analysis at design (e.g. in Rwanda, Senegal and São Tomé and Príncipe) were based on previous experience in a given area and a set of commodities. From an initial focus on increasing production and productivity, these projects had transitioned to supporting producers' access to market, processing and retailing facilities.

Few project designs were backed by a systemic perspective on value chains encompassing market characteristics, opportunities and trends; price evolution over time and locations; or estimation of initial investments and costs for small-scale producers.

Projects have sought to help small-scale producers and other value chain stakeholders manage production-related risks by providing training on improved agronomic practices and control of pests and diseases. Logistical and infrastructure-related risks have been addressed by constructing or rehabilitating rural roads and bridges. Projects had less focus on market and price risks than on infrastructure. An example was the price crash in the raspberry value chain in Bosnia and Herzegovina, which was not anticipated, although it was known that the country was a small producer surrounded by large producing countries and that prices would be profitable for small producers when the neighbour countries experienced low harvests.

Most projects did not address policy and enabling environment challenges and risks, although there were also exceptions, such as in Sudan (gum Arabic value chain). There, cofinancing with the World Bank helped turn a national purchasing board authority, which kept farm-gate prices low, into a regulatory authority and opened the market to private traders, leading to higher prices to producers. In Kenya, one project worked on the regulation of the horticulture subsector and another on policies for the dairy subsector. Regulation on and verification of product standards, labelling and food safety are likely to become a priority for international and domestic markets.

AfDB IDEV (2018) found that lack of full value chain analyses and market studies have limited the relevance of its operations. Each value chain intervention is expected to ensure added value along the chain for as many actors as possible, without which other actors may not support improvement in one link of the chain, which might adversely affect the achievement of outcomes. However, the country case studies found that, in practice, few interventions involved a systemic analysis to ensure that the interventions were relevant. For example, in the Democratic Republic of Congo (DRC), insufficient consideration was given to equipment for facilities constructed

to enable viable operations for meat value chains. In Zambia, there was a focus on increasing cashew production and infrastructure for processing but insufficient analysis of the interplay between the international and domestic markets and how increased production would be absorbed in the markets.

Flexibility in responding to market changes was not adequately considered in the design and implementation modalities. Value chain development interventions cannot be planned fully in advance of an intervention. During the course of implementation, market factors and actors may change (e.g. export price fluctuations for cocoa in Côte d'Ivoire and cashews in Zambia). Therefore, it is critical that implementers of interventions have the capacity to respond to market signals and review the original analyses to assess whether they are still relevant. Adaptation to changing contexts calls for a robust monitoring and evaluation system and room for adaptive management in project design that allows projects to be responsive to changes in the value chain context or markets for the targeted commodities. The country case studies illustrated that there was insufficient monitoring and evaluation to assess the extent of impact and sustainability. During implementation, lack of consideration of responsiveness to market needs caused sustainability challenges (AfDB IDEV 2018).

Approaches That Projects Took to Support Inclusive Value Chain Development

IFAD-funded projects took various approaches to value chain development (table 13.1). Products and processes were upgraded, and horizontal linkages, which were derivative of IFAD's traditional project approaches, were strengthened in the vast majority of projects. This suggests that production aspects required improvement before interventions could strengthen vertical linkages or functional upgrading, which were seldom observed. This may also indicate lack of clarity regarding how to facilitate access to the three value chain flows – commodity, money and information – to maximize their benefits in the process³.

³ *Product upgrading* is an increase in the quality or quantity of production (production techniques, higher-value products). *Process upgrading* is an increase in the efficiency of the production process to reduce production costs and promote certification, food safety or traceability. *Strengthening horizontal linkages* refers to improving linkages among stakeholders at the same functional level of the value chain (e.g. creation of cooperatives, federations, capacity building of producer organizations) to increase their bargaining power to buy their inputs and sell their

Table 13.1 Number and Percentage of Reviewed Projects That Included Different Aspects of Value Chain Strengthening in Design

Value chain segments addressed	Number	Percent
Product and process upgrading	75	97.4
Horizontal linkages	67	87.0
Vertical linkages	61	79.2
Governance mechanisms	51	66.2
Functional upgrading	44	57.1
Enabling policy environment	28	36.3
Market information systems	11	14.3

Source: IFAD IOE (2019).

Note: $n = 77$.

Market information systems were planned in only 14 per cent of projects reviewed at IFAD. The main challenges had to do with the time required to establish market information systems and to ensure that these systems were institutionalized and financially sustainable after the end of project funding.

AfDB's support tended to focus on the primary production segment, with the greatest proportion of resources dedicated to infrastructure, equipment and inputs in support of production (irrigation, seeds and seedlings) and to a lesser extent on processing and marketing (bulking centres, landing sites, milk collection centres and market sheds). Few projects strengthened links between actors (public, private, farmer's organizations, civil society) or fostered agreements between them (contracts and trust building). Across the nine case study countries, although production was supported in some way in all nine commodities studied, value addition was supported in only six (Zambia, Rwanda, DRC, Liberia, Uganda, Mozambique). These mainly

outputs. *Strengthening vertical linkages* means improving linkages among stakeholders at different functional levels of the value chain. This may include promoting formal or stable types of contracting and increasing physical access to markets. *Functional upgrading* refers to adding new functions and activities to the target group (e.g. producers and their associations), such as processing, storage and packaging, to capture more value (IFAD IOE 2019).

related to provision or rehabilitation of market infrastructure and processing units and some training and extension for commercialization⁴.

Public sector support can enhance value chain development but requires good working relationships with private sector actors and other relevant organizations such as farmers associations, rural banks and input suppliers. Examples of public–private collaboration were, on the one hand, the success in Rwanda with milk collection centres and processors’ and farmers’ associations and, on the other hand, challenges with credit access in Mozambique, where linkages were not effectively established, thwarting the desired outcomes.

Many interventions lacked private sector engagement and market orientation. For instance, the Rural Infrastructure Development Support project in DRC was designed in 2010 as a rural infrastructure project, and in Mozambique, the Baixa Limpopo Irrigation and Climate Resilience Project invested in irrigation infrastructure and did not specifically aim to support rice marketing. In neither of these cases were the facilities constructed used to their full potential because market factors were insufficiently considered. This could have been addressed early on in the projects if private sector actors had been engaged in determining market needs and size. A positive example was the dairy farmers’ cooperatives in Rwanda, which gave collective voice to dairy farmers, generated economies of scale, enhanced product quality and engaged in marketing on behalf of farmers (AfDB IDEV 2018).

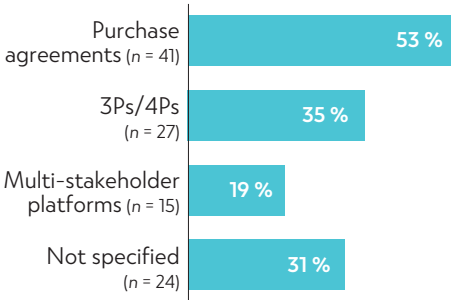
Making Governance of Value Chains More Inclusive

Mechanisms to improve value chain governance were promoted in two-thirds of the projects reviewed (IFAD IOE 2019). Purchase agreements between producers and buyers were the most common form of governance, involving 53 per cent of projects, with 35 per cent promoting public–private–producer partnership arrangements and 19 per cent supporting multi-stakeholder platforms. (Approximately one-third had no governance arrangement (figure 13.3).

Purchase agreements ranged from loose, informal agreements to fully defined contracts that specified the quantity, quality and price of goods and

⁴ For example, support to the meat subsector in DRC included rehabilitation of slaughter facilities and markets, and support to the cassava value chain in Liberia included processing and training for commercialization. In addition, because of the lack of value chain analysis, the profitability of value-adding activities was not clearly defined in AfDB’s interventions.

Figure 13.3 Governance Mechanisms Used in Projects Reviewed



Source: IFAD IOE (2019).

Note: $n = 77$. Projects may have more than one governance mechanism, so numbers do not sum to 100 per cent. 3P = public-private partnership; 4P = public-private-producer partnership.

the terms of the transaction. Some projects facilitated agreements between producer groups and processors, for example, the rice value chain in Cambodia. Other projects enabled producer organizations to better supply clients according to precise requirements for quality and delivery (e.g. palm oil bunch in Uganda; coffee, cocoa, cashew and horticulture cooperatives in El Salvador and Honduras).

Public-private-producer partnerships are agreements between government agencies, private sector entrepreneurs and producer organizations. They were instrumental in motivating private sector engage-

ment in pro-poor value chains, although many interventions did not address fundamental questions regarding incentives for entrepreneurs to partner with small-scale producers and requirements such as the size of the initial investment (training, machinery), the expected profit margin and risks, and the size of the market and level of competition.

Nineteen per cent of projects reviewed set out to form multi-stakeholder platforms, which bring together stakeholders linked to a value chain (e.g. input providers, producers, processors, distributors) to increase communication, trust and mutual understanding and establish commercial relationships. Establishing these platforms was an advanced way to improve governance of the value chain. This functioned well where there was a tradition of dialogue among stakeholders, such as in Niger and Senegal, but the role of projects in enabling all actors to participate actively was equally important.

More far-reaching results in terms of changes in governance were found in the projects in which multi-stakeholder platforms had been established and worked well (e.g. Nepal, Niger, Senegal and, in part, Ghana and Uganda). The platforms opened space for dialogue and coordination regarding issues such as input supply, market infrastructure, price level, market information and dispute resolution.

Value chains that straddle countries, such as with tradable cash crops (e.g. coffee, cocoa, cashews, dried fruits), are a special challenge. Key value

chain stakeholders are located in a different country from the one where the development project is supported and may be difficult to reach. Moreover, trade policies in other countries may affect demand for and prices of products. As the IFAD evaluation found, a generally successful way to address these challenges was to link producer organizations with fair-trade movements. This helped these organizations negotiate special price premia related to production modality (e.g. organic, low-chemical inputs; good stewardship practices for natural resources) and bridge the gap with stakeholders in other countries. Reportedly, producer organizations linked to fair-trade movements experienced smaller fluctuations in commodity prices, although this required long-term support of producer organizations (e.g. extension, quality control), as well as policy and regulatory frameworks of national governments (e.g. inspection for sanitary and phytosanitary standards), and was hard to achieve during a single project phase (6–7 years). Successful cases were found when there were two or more coordinated project phases.

Another area that requires attention is trade policy and regulatory coordination between countries, because projects are typically focused on individual countries. Some initial attempts to address bilateral trade coordination were found between Niger and Nigeria, promoted by an IFAD-funded project in Niger. Similarly, AfDB supported regulatory policy coordination between Uganda and DRC in fish resource management and marketing. In Côte d'Ivoire, AfDB also increased the capacity of producer organizations, which resulted in better quality cocoa because obtaining certification required that international standards be met.

Evidence of the distribution of value within value chains was fragmented, but the distribution appeared to be more stable and equitable when efforts were made to develop dialogue and trust between stakeholders, producer organizations were empowered to negotiate exchange conditions, competition was high between buyers (so that they had to offer good prices and other favourable transaction conditions to attract small producers), focus was on niche markets and buyers were committed to fair terms of trade.

Financing the Value Chain

Projects were effective at providing basic financial services to producers through community-level informal groups and some microfinance institutions (IFAD IOE 2019), although projects offered conventional rural finance services rather than instruments specific to value chain financing. The most

common instruments were linkage facilitation between formal and informal financial institutions; credit that rural finance institutions provided to small-scale producers, generally short-term finance for purchasing inputs; matching grants for small-scale producers to reduce the total amount borrowed and grants to aggregators, processors and wholesalers to offset costs and encourage partnerships with small-scale producers and their associations.

Experience in financing small and medium-sized enterprises, cooperatives and producer organizations was uneven. In turn, these organizations could not offer prompt cash payment to their members, creating incentives for side selling and making it difficult to fulfil purchase agreements with buyers. Part of the problem was the lack of familiarity of banks with the specific agribusiness finance systems and hence their aversion to offering agricultural credit. From the borrower's side, cooperatives and producer organizations faced small profit margins and could not afford prevailing interest rates.

In five of the nine cases, there was a variety of financial intermediation support (AfDB IDEV 2018). For instance, in Rwanda, support was provided to dairy cooperatives to access finance to support members. In DRC, a project coordinated with a microfinance institution to increase access to finance in the project area. In Morocco, a project helped farmers access agricultural insurance, which increased their access to formal credit. In Rwanda, to reduce risk of financial losses, finance was provided in kind (one cow per low-income family); through this scheme, 16,072 families received cows, with repayment deducted through the cooperatives, which helped increase dairy production in the country by 59.6 per cent and helped reduce poverty (from 44.9 per cent to 39.1 per cent) in beneficiary families. In Mozambique, access to credit was limited, which severely restricted benefits to poor farmers.

Promoting Transformative Changes for the Poor

As noted, the shift towards value chain support was expected to promote transformative changes for the poor, although there has been debate, at IFAD, AfDB and elsewhere, whether it is feasible to reach out to very poor groups through value chain approaches. Based on the evaluation findings, a short (perhaps crude) answer is that it is possible but will not happen automatically and requires a clear sense of direction and good diagnostics – at project design and during implementation. Factors contributing to effective outreach to poor small-scale producers included (IFAD IOE 2019):

- selecting commodities requiring little land or capital investment and involving intensive, unskilled labour inputs;
- enforcing pro-poor requirements for agribusinesses as a condition for obtaining IFAD project support;
- community-based groundwork and mobilization of producer groups combined with other activities; and
- previous work in the same area establishing the productive base and local knowledge and participatory approach to design and implementation.

Targeting was often weak when there were unwarranted assumptions about trickle-down effects to poorer groups from more entrepreneurial farmers and agribusinesses. Such effects might take place when there was a sizeable increase in demand for smallholders' products and a significant increase in farm-gate prices (e.g. Vietnamese coconut processing) or sizeable effects on demand for unskilled or semiskilled labour (e.g. in El Salvador, Honduras and Rwanda). In many cases, assumptions regarding trickle-down effects had not been appraised *ex ante* and did not materialize.

In terms of gender equality, better results were achieved in projects that selected value chains involving large numbers of women as producers or processors (e.g. food crops, small ruminants, artisanal products, agro-processing). A crucial factor was how structural causes of gender inequalities, including social norms and distribution of economic resources at all levels of the value chain, were addressed.

AfDB IDEV (2018) devised strategies to enhance inclusiveness in value chain development. The portfolio review found that 63 per cent of interventions assessed had design elements to address inclusiveness in terms of gender, youth or other vulnerable groups. The review found that more-recent designs using a value chain development approach linked clearly vulnerable groups to markets, as in the Malawi Agricultural Infrastructure and Youth Agribusiness Project. This project supported 'youth entrepreneurship, storage agro-processing and value addition through market linkages and trade facilitation, linking farmers with agro-processors, building bulk commodity network, eliminating middlemen, and exerting group effects on processors for better prices' (AfDB IDEV 2018, 37).

AfDB IDEV (2018) concluded that its processes lacked a systemic approach to inclusiveness, examining not only the position of vulnerable populations all along the value chain, but also their capacity to access productive assets (water, capital, knowledge, land), low literacy levels, lack of formal representation and the social norms they encounter within their communities

and households. More specifically, in Mozambique, quotas have ensured that vulnerable populations attend capacity-building sessions, but no additional measures were developed to ensure anything beyond participation. Similarly, in Liberia, quotas have ensured that vulnerable populations receive training and cassava cuttings, but the benefits do not seem to extend much further.

Evaluative evidence of value chain development at AfDB confirmed the importance of gender-sensitive analysis at design and throughout implementation, as well as of preparing and implementing gender action plans to ensure that intended impacts reach women and adverse consequences are avoided. Evidence from case studies shows that some projects have developed gender plans (e.g. Zambia and Liberia). In Zambia cashew value chains, gender has been considered during the planning process, mainly through allocation of quotas (50 per cent of training session participants to be women), but it was not clear whether participation would be sufficient to include equal benefits for participating women according to poverty level or vulnerability. Inclusiveness was not sufficiently budgeted for or integrated into implementation and monitoring and evaluation mechanisms. It was not clear from the portfolio review how the analysis of gender and youth issues included in the design were managed during implementation of projects and programmes.

AfDB country case studies have shown that ensuring participation of more-vulnerable segments of the population in project activities (by assigning quotas) is necessary but insufficient to ensure that they benefit proportionally. In large infrastructure projects in Mozambique, Morocco and DRC, gender and other inclusiveness-related factors were not tracked, making it difficult to ascertain whether benefits had reached vulnerable target groups.

Pathways Towards Transformative Changes for Poverty Reduction

Despite significant variations between countries and projects, there were many examples of considerable increases in productivity, combined with better access to markets and timing of marketing, higher farm-gate prices and greater diversification of marketed products with good sustainability prospects with little external support (IFAD IOE 2019).

The mechanisms through which value chain participation could be transformative for the poor included:

- improvements in product characteristics (e.g. larger, better-looking fruit in Morocco) or a shift to higher-value products (e.g. vegetable crops or fruits in China);

- price mechanisms, such as ex ante agreement on a fixed price to reduce risks of price fluctuation for producers and price premia linked to product characteristics (e.g. organically grown coconuts in Vietnam);
- improvements in producers' capacity to negotiate output prices and greater economies of scale for producers, thanks to horizontal linkages (e.g. in Honduras and El Salvador);
- capturing value added through functional upgrading (e.g. through processing and reducing the role of middlemen); and
- employment generation – for which evidence was incomplete, although in some value chains, such as coffee, horticulture and dairy (e.g. Bosnia and Herzegovina, El Salvador, Honduras, Rwanda), the evaluation observed greater use of waged labour in producer organizations and agribusinesses stemming from project interventions.

Mapping the Emerging Findings

IFAD IOE (2019) mapped a number of value chains supported by projects, along with two main indicators: level of development of value chains (incipient, intermediate, advanced)⁵ and degree to which value chains were generating pro-poor outcomes (low, medium, high)⁶. With regard to

⁵ *Incipient* value chains were defined as those that involve the primary steps of mobilizing small-scale producers, providing training on productivity and quality, increasing access to inputs and production credit and building feeder roads and simple market infrastructure for greater market access. For *intermediate* value chains, priorities were organizational strengthening and functional upgrading for producer organizations, early development of vertical linkages, financial resources for value chain infrastructure and technology (e.g. warehouses, cold stores, processing machinery) and organized marketing of products. *Advanced* value chains involved a higher level of product; process and functional upgrading (e.g. through certification or branding); more-specialized technical assistance and capacity building (including on financial literacy and business management); finance for investment and working capital; development of purchase agreements with buyers; some form of risk management and market information systems; and structured dialogue among value chain stakeholders, including government bodies, for example, through multi-stakeholder platforms.

⁶ Four criteria were used to categorize the degree of pro-poor outcomes: inclusiveness (degree of actual poverty outreach), empowerment of people and groups, size of benefits for the poor (e.g. income, food security) and perspectives for sustainability of benefits for the poor. Value chains considered strong on all of these criteria were categorized as *high* in terms of pro-poor outcomes, those that were strong on only two criteria or for which performance was reasonably good on all four criteria were rated as *medium*, and those with poor performance on most criteria were categorized as *low*.

value chain development, 35 per cent of cases were incipient, 41 per cent intermediate and 23 per cent advanced. In terms of pro-poor outcomes, 33 per cent were low, 44 per cent medium and 22 per cent high (table 13.2), an overall favourable finding.

Table 13.2 Mapping of Projects and Value Chains According to Level of Development and Pro-Poor Outcomes (Percentage of Observations)

Value chain development level	Pro-poor outcomes		
	Low	Medium	High
Advanced	3	10	10
Intermediate	10	19	12
Incipient	20	15	0

Source: IFAD IOE (2019).

In the 20 per cent of projects that did not have clearly articulated value chain designs and whose implementation did not go beyond supporting production, value chains were found to be incipient and not achieve pro-poor outcomes (table 13.2). At the other end of the spectrum, 10 per cent of the value chains reviewed reached an advanced development stage and achieved pro-poor outcomes, making a powerful case for the value of the projects. It can be argued that, in such cases, interventions had been transformative for value chains and for poor, small-scale producers. A common trait of these transformative interventions was that IFAD had long experience in the project area and had supported multi-stakeholder platforms and interprofessional associations (value chain governance) and had benefited from specialized technical assistance to support the project management team.

Key Conclusions and Policy Implications

The AfDB and IFAD evaluations both concluded that it is worth investing in support for inclusive value chains, which can have transformative effects on poverty reduction and development in rural areas, although such support is conceptually complex and requires a systemic perspective and transformation in the capacity, skills and organization of the supporting agencies.

The AfDB evaluation identified five fundamentals to be applied in all value chain interventions (table 13.3): careful context-specific value chain analysis to ensure addition of value along the chain; inclusion of poor farmers, women, youth and other vulnerable groups in participation and benefit sharing; flexibility and responsiveness to changing contexts and market needs; a primary focus on the profitability and efficiency of the value chain; and application of strategies to ensure sustainability of outcomes.

It also recognized five enabling factors (table 13.3) that are more context specific than the five fundamentals and are good predictors of positive outcomes in developing pro-poor value chains: appropriate infrastructure and technology, policy and regulatory environment favourable to the targeted value chain, appropriate business support services to improve the skills of value chain actors, access to finance for value chain actors to make necessary investments to increase profitability, and private sector engagement and working relationships between value chain actors.

Table 13.3 Fundamental Factors and Key Enablers for Value Chain Development Interventions

Factor	Key component from a systemic perspective
Fundamentals	
Value chain analysis	<ul style="list-style-type: none"> ▪ Constraints of the value chain ▪ Understanding the socioeconomic factors of the target group ▪ Stakeholder mapping and power relationship in the value chain ▪ Value added distribution ▪ Potential market ▪ Risk assessment and mitigations
Profitability with value addition	Financial and economic viability of added values in the value chain
Responsiveness to market	<ul style="list-style-type: none"> ▪ Ability to respond and adapt to market requirements to secure business in the face of competition ▪ Monitoring and evaluation system linked to value chain
Inclusiveness	<ul style="list-style-type: none"> ▪ Involvement of women, youth and the poor in value chain ▪ Attention given to women, youth and the poor in planning and implementation of interventions ▪ Evidence of benefits to women, youth and the poor

(continued)

Table 13.3 Fundamental Factors and Key Enablers for Value Chain Development Interventions (*continued*)

Factor	Key component from a systemic perspective
Sustained impact	<ul style="list-style-type: none"> ▪ Technical ▪ Financial, economic ▪ Institutional ▪ Political, sociocultural
Key enablers	
Infrastructure and technology	<ul style="list-style-type: none"> ▪ Irrigation, access roads, market sheds, storage houses, processing units ▪ Improved inputs (seeds, fertilizers, agricultural tools) ▪ Information and communication technology
Policy and regulatory (business) environment	<ul style="list-style-type: none"> ▪ Rules and regulation to improve business environment ▪ Policy dialogue to improve value chain structure and governance ▪ Activities to improve quality standards
Access to finance	<ul style="list-style-type: none"> ▪ Credit facility (in cash or kind) ▪ Contract farming ▪ Risk-sharing facilities ▪ Cascade financing schemes from distributors to processors and producers
Business support	<ul style="list-style-type: none"> ▪ Organizational capacity ▪ Market access support ▪ Entrepreneurial skills such as financial analysis and management, process monitoring and management, and human resource management ▪ Technical skills
Private sector participation and linkages among value chain actors	<ul style="list-style-type: none"> ▪ Private sector engagement ▪ Collaboration among value chain actors ▪ Trust building in value chain ▪ Information management

Source: AfDB IDEV (2018).

The two evaluations provided recommendations on how international development organizations could better support value chain development. Some common elements were the following.

First, organizations need conceptual clarity on what a value chain is and what the critical requirements are to make them viable, sustainable and inclusive. A systemic perspective on value chains, such as the one

presented in this chapter, can help provide clarity. Organizations need to ensure that they have the internal capacity and resources to design, supervise and support the execution of programmes promoting value chain development. They need to ensure that key partners in the countries have adequate skills and experience in value chain support. Government entities, non-governmental organizations and even private entrepreneurs may need special support, for example through targeted technical assistance initiatives. In many organizations, a strategy or an action plan would help provide a more coherent, ideally system-based, approach.

Second, value chain development requires long-term engagement. In many financial institutions, this often entails providing support throughout several project phases. At the design stage, projects cannot assume that value chain support is what is needed. Instead, they should systematically assess the degree of preparedness for value chain support, taking into account the local context and previous experience of the government and the funding organization. If value chain support is not the right starting point, a more conventional approach, such as supporting production or transportation infrastructure, may be the first step to take. A value chain approach may be adopted later, for example in the next project-financing phase.

Third, projects need to actively promote gender equality and outreach to poor and very poor groups rather than assuming that trickle-down mechanisms will be operating, like an 'invisible hand'. Project designs should provide a theory of change explaining how benefits reach very poor groups (e.g. through wage employment generation or higher demand and higher farm-gate prices) and identify major barriers and how to overcome them. These assumptions must be corroborated by evidence.

Fourth, projects need to promote inclusive value chain governance and an inclusive policy and regulatory environment by establishing or strengthening multi-stakeholder platforms and interprofessional associations that provide small-scale producers and other value chain stakeholders with information on prices and markets, a venue for dispute resolution and a voice in discussing the policy and regulatory system.

Fifth, in addition to conventional approaches to rural finance (e.g. linkage of banks with village-level groups, credit lines and matching grants for individual borrowers), there is a need to devise ways to provide financing along the value chain, including producers, buyers, processors and retailers. Typically, small and medium-sized enterprises, cooperatives and companies active in aggregation of produce, transformation and distribution had limited access to finance at an affordable interest rate. This generated

cash-flow problems and constrained their capacity to procure from small producers. Whole value chain financing schemes (cascade financing from distributors to processors to producers) are used in industrialized countries and are emerging in developing countries.

Sixth, no single organization can manage complex endeavours alone. Development agencies need to strengthen partnerships with other organizations, including private sector organizations that have value chain expertise, to ensure that projects are based on a thorough analysis of commodity market structure, demand and supply, price level and volatility, and barriers that small-scale producers face.

Implications for Evaluations of Value Chain Support for Poverty Reduction

What are the lessons from these evaluations for future evaluative work on value chains? The two evaluations discussed in this chapter were conducted at the corporate level but also tracked the results that AfDB and IFAD achieved through project-funded operations.

Having been conducted at the corporate level, the evaluations assessed to what extent the organizational structure, such as preparation of specific strategies, recruitment of specialized staff, adoption of technical guidelines, training of staff and capacity building of government implementation units, had changed to accompany the increase in focus on value chain development. The analysis concluded that the pace of organizational change had not matched the shift in attention being paid to value chains in the lending portfolio. Future evaluations of this type with a thematic or corporate scope should review institutional capacity to support value chain development. As argued in this chapter, for an institution to be transformative requires that it transform itself.

Tools for conducting the institutional analysis may include reviews of an organization's documentation (corporate strategies, country strategies, project report, organizational charts, specific budget and human resources allocated to quality assurance, technical support to value chains); review of operational and organizational experience of peer organizations; interviews with executive board members, organization managers, and staff and development counterparts in the country where projects are supported; and an electronic survey of the organization's staff and project managers to determine their knowledge, views and experience.

An opportunity that AfDB and IFAD did not pursue, but that deserves consideration, is that of conducting joint evaluations, particularly when

there are cofunded or cosponsored initiatives and projects. Joint evaluations can be challenging to manage as the numbers of decision makers and stakeholders increase. At the same time, they allow for organizational cross-learning and may deliver a stronger message to senior management and governing bodies.

The AfDB and IFAD evaluations also attempted to assess the effects of value chain development projects on household and community welfare. The two evaluations identified pathways and conditions through which engagement in value chains could become transformative for the lives of small rural producers. They found uneven evidence as to whether projects had been transformative. Part of the problem was that value chain projects belonged to more recent project cohorts, many of which were ongoing, so the full range of effects was not yet discernible. Another problem was the absence of well-established data (e.g. from surveys). Many future evaluations, even if conducted at the project level, are likely to encounter the same constraints. It might be sensible for many of them to manage their ambitions, concentrate on identifying the pathways through which a value chain project could produce transformative results and assess whether the project has developed those pathways. When the budget and time to conduct an evaluation are limited, this could be more effective than designing complex, time-consuming surveys.

Is a systemic conceptual framework useful for conducting an evaluation on a value chain-related topic? The conceptualization of value chains as systems underpinned the evaluations reviewed in this chapter. This is useful for understanding the complexity of developing an intervention for inclusive value chain development and the interconnectedness of a value chain system. The systemic approach can be a good conceptual reference even for project-level evaluations. Although a project supporting value chain development may concentrate on only one subsystem or a node within that subsystem (e.g. market infrastructure, processing of raw products), the evaluator would still benefit from awareness of the bigger picture that a systemic approach provides. This would help explain the importance of other subsystems or nodes of subsystems (e.g. governance of the value chain, the policy, the regulatory system). In general, a good approach should be system aware, even if not system centred.

Note

The opinions expressed here are those of the authors only and do not represent the official position of the organizations with which they are affiliated.

References

- AfDB IDEV (African Development Bank Independent Development Evaluation). 2018. *The African Development Bank's support for agricultural value chains development: Lessons for the Feed Africa Strategy*. Abidjan, Côte d'Ivoire: AfDB.
- Arias, Pedro, David Hallam, Ekaterina Krivonos et al. 2013. *Smallholder integration in changing food markets*. Rome: Food and Agriculture Organization of the United Nations.
- FAO (Food and Agriculture Organization of the United Nations). 2014. *Developing sustainable value food value chains – guiding principles*. Rome: FAO.
- IFAD IOE (International Fund for Agricultural Development Independent Office of Evaluation). 2019. *Corporate-level evaluation on IFAD's support to pro-poor value chain development*. Rome: IFAD.
- Kaplinsky, Raphael, and Mike Morris. 2002. *A handbook for value chain research*. Brighton, UK: Institute of Development Studies.
- M4P (Making Markets Work for the Poor) and DFID (UK Department for International Development). 2008. *Making value chains work better for the poor. A toolkit for practitioners of value chain analysis*. Phnom Penh, Cambodia: Agricultural Development International.
- Porter, Michael E. 1985. *The competitive advantage: Creating and sustaining superior performance*. New York: Free Press.
- USAID (United States Agency for International Development). 2014. *A framework for inclusive market system development*. Washington, DC: USAID.



CHAPTER 14

Evaluating Under Fragility: Lessons from the Palestinian Context

KHALED RAJAB

Abstract. Like that of other countries, the Palestinian government committed itself to realizing a number of Sustainable Development Goals (SDGs) and has put in place a national mechanism for tracking and monitoring progress towards achieving them, including establishing institutional frameworks; engaging different stakeholders, including civil society, donors and the private sector; identifying indicators; collecting data and producing the Voluntary National Review Report. Nevertheless, the unique context of Palestine as a fragile, conflict-affected country poses a number of challenges, as it affects the ability of the national government and other societal actors to monitor achievement of the SDGs, to say nothing of actually achieving these goals. This chapter contributes to the discussion of the challenges of monitoring and evaluating SDGs in fragile contexts by focusing on the experience of Palestine. It builds on findings from literature review and interviews with relevant stakeholders, including government, civil society and international development partners that support Palestine in this effort, in particular the various United Nations agencies. It also presents and discusses key lessons from the Palestinian experience, especially for other countries under similar circumstances, including in the Middle East and North Africa.

Introduction

Five years have passed since all members of the United Nations General Assembly ratified the 2030 Agenda for Sustainable Development and its core, the Sustainable Development Goals (SDGs). Comprising 17 goals and 165 targets, the SDGs represent a partnership between the developed and developing world to address complex and interdependent challenges in today's globalized world. These goals strive to address critical challenges facing the world today, including eradicating extreme poverty, global inequality and climate change; promoting sustainable urbanization and industrial development; protecting natural ecosystems and fostering growth of peaceful, inclusive communities and governing institutions (Chaitanya Kanuri 2016). The goals provide clear guidelines and targets for all countries to adopt in accordance with their own priorities and challenges. Not all countries are equal in terms of their ability and readiness to achieve the SDGs, and many will have to address their own, often distinctive challenges, requiring flexibility and innovation in how to translate the global SDGs into realistic, locally adapted policies and interventions. Fragile states are among those that necessitate unorthodox approaches and instruments to fulfil their commitment to achieving the SDGs by 2030.

The author presented and discussed an outline of this chapter in a special panel at the IDEAS Global Assembly, October 2019, Prague, Czech Republic. The stimulating discussion during that panel indicated a need for further research on the topic of fragile states and how they address the challenges of implementing and tracking the SDGs. This chapter contributes to this discussion by presenting the experience of Palestine in tracking and monitoring the SDGs and hopes to draw the attention of the evaluation community and development practitioners to a more utilization-focused evaluation¹ that will enhance our understanding of fragile contexts and lead to relevant, effective policies and interventions.

To produce this chapter, the author relied on information from a literature review and interviews with relevant stakeholders, including the government, statistical bureau, civil society and international development partners who support Palestine in fulfilling its commitment to the SDGs.

¹ Utilization-focused evaluation is an approach based on the principle that an evaluation should be judged on its usefulness to its intended users (Patton and Horton 2009).

Theoretical Framework of Fragility and the SDGs

I start by offering a theoretical background on fragility and the SDGs to frame our understanding of the nature and challenges of fragility before delving deeper into Palestine's experience as a fragile state in achieving, monitoring and evaluating the SDGs.

The concepts of fragile states and failed states have been of practical interest to academics, policymakers, the evaluation community and international multilateral and unilateral organizations since they were developed (François and Sud 2006). Although there is no agreed-upon definition of a fragile state, most development agencies and practitioners apply the term to a situation in which the central government fails or is unable to perform its core functions of ensuring basic security, maintaining rule of law and justice and providing basic services and economic opportunities for its people, including the poor (McCloughlin 2010). The literature provides various definitions of fragile state (e.g. Eizenstat, Porter and Weinstein 2005; François and Sud 2006; Newbrander 2012). Two critical elements can be distilled from the various definitions: lack of legitimacy and lack of effectiveness in providing security and services (Newbrander 2012). Other scholars link fragility with capacity deficits (Brinkerhoff 2010) and inability to protect essential civil freedoms (Eizenstat, Porter and Weinstein 2005). Hagesteijn (2008) describes fragile states as having weak institutional capacity, limited control of land and territory and inability to fulfil their mandate to provide services to and ensure the welfare of their constituents.

This inability of a government to meet citizens' expectations exacerbates problems of legitimacy and effectiveness (Brinkerhoff 2005). It has been argued that one of the defining characteristics of fragility is a weak state of legitimacy, under which fragile states 'fail to establish reciprocal state-society relations or create a binding social contract' (McCloughlin 2010, 5). Moreover, the inability of a government to deliver core services (e.g. transport, electricity, health, education, water, sanitation) to its citizens and to provide a decent level of economic opportunity and welfare calls into question the legitimacy of that government. It also results in citizens losing trust in national institutions and withdrawing their support, jeopardizing the nation-building process (Brinkerhoff 2005; Rakodi 2001).

The connection between state fragility, legitimacy and state building is well documented in the academic and development literature (Brinkerhoff 2010; Lister 2005; Roberts 1990). The term 'state building' refers to the process of creating a functioning state and was historically used in the context of constructing Western European states (Sekhar 2010). The term

was then expanded to fragile states with the aim of helping them improve quality of life for citizens and establish policies, institutions and governance arrangements to support socioeconomic development (Brinkerhoff 2010) and citizen welfare (Roberts 1990). Brinkerhoff (2005) argues that building the state's capacity strengthens its legitimacy and prevents the risks of backsliding. Moreover, it is crucial for the international community to have a better understanding of capacity and capacity development (Brinkerhoff 2010) and concentrate on strengthening the domestic capacity of the state and not just promote democracy (François and Sud 2006). This is important particularly in the context of fragile states where organizations are being established by external interventions. Then again, transforming these organizations into legitimate institutions requires time and depends on domestic political processes (Ottaway 2002). International aid organizations should avoid bypassing the government by using other delivery mechanisms, which will damage pre-existing capacity and undermine state capacity-building efforts (François and Sud 2006).

The number of fragile states increased from 56 in 2016 to 58 in 2018, with 15 classified as extremely fragile and 43 as fragile (OECD 2018). Most of the 58 fragile countries are facing great challenges in ending extreme poverty². Although the percentage of people living in extreme poverty globally has decreased significantly, from 28 per cent in 1999 to 11 per cent in 2013 (UN DESA 2017), the percentage of people living in extreme poverty is projected to rise in 40 of the 58 fragile states by 2030. Moreover, with the assumption that no action will be taken, the Organisation for Economic Co-operation and Development (OECD) estimates that the number of people living below the international poverty line in fragile situations will increase from 513.6 million in 2015 to 620 million in 2030. In other words, more than 80 per cent of the world's poorest people could be living in fragile contexts by 2030 (OECD 2018).

In 2018, Palestine³ witnessed only slight improvements in the societal and environmental dimensions of fragility (OECD 2018), with the improvement in aggregate fragility mostly from the security dimension because the risk of violent conflict has decreased in the past few years. Despite Palestine's upward trend in the fragility framework and improvement in overall security, it still ranks second worst in the world, just after Syria, in control

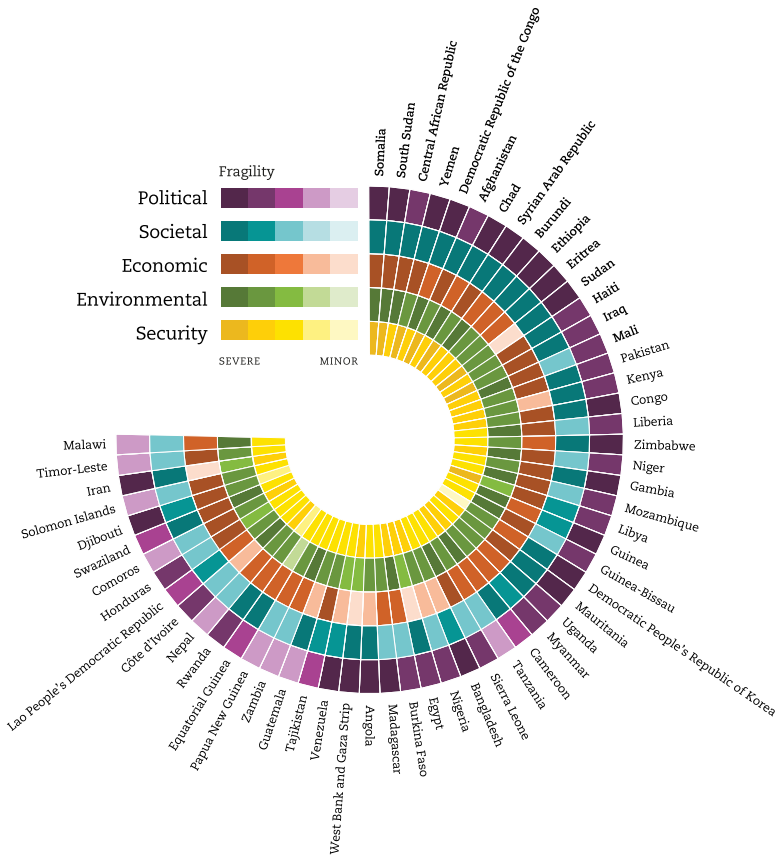
² The international community defines extreme poverty as living on less than \$1.90 a day, as measured in 2011 international prices (equivalent to \$2.12 in 2018) (Marcio Cruz 2015).

³ OECD used the term 'West Bank and Gaza' instead of 'Palestine'.

over territory. In addition, the political dimension has worsened because of a decrease in voice and accountability and an increase in political terror (OECD 2018).

Figure 14.1 illustrates the state-of-fragility framework, showing the West Bank and Gaza (or Palestine) among the countries that face severe fragility in various dimensions.

Figure 14.1 State-of-Fragility Framework, 2018



Source: OECD (2018), p. 83.

SDGs in Fragile Contexts

Despite receiving praise from many researchers and development practitioners, the SDGs have not escaped criticism. For example, they were criticized because of their similarities to the post-2015 development agenda, which includes the Millennium Development Goals (MDGs) and the post-MDG processes. SDGs are no different from the MDGs in terms of having a donor-driven design that led to neglect of problems in developing countries and lack of consideration of real needs of recipient countries, particularly those of marginalized populations (Miyazawa 2012) and particularly in fragile contexts.

The challenge of advancing the sustainable agenda in fragile states was noted in the academic and development literature. Carment (2017) argues that, given historical evidence of the lack of progress on the MDGs in countries affected by conflict and fragility, achieving the SDGs will be a challenge. Countries affected by conflict and fragility often have the most limited capacities and resources while facing the greatest political challenges. The multifaceted challenges that fragile states face require that the decision-making process address not only technical, but also political obstacles to development (Carius 2018).

It is crucial to focus on developing state institutions in fragile states, which is a very long-term process, often taking decades (OECD 2014). In the context of SDGs, fragile states should focus on long-term rather than short-term results. For instance, working in fragile states, and with marginalized groups in more stable contexts, will take time and involves risk. Therefore, fragile states often need to develop institutional capacity: a very long-term process, often taking decades (Greenhill 2016).

Other researchers have noted the difficult task of measuring the SDGs because there are many SDGs and subgoals and few real measurements or baselines, priorities or even clarity of basic definitions of such terms as 'sustainable' and 'development' (Lempert 2017). Those with this critical view claim that, in general, the SDGs present little change in substantive, ideological or implementation approach from the MDGs that would offset the deficiencies of the MDGs. According to this criticism, the SDGs are still promoting an agenda of globalization, urbanization and assimilation that does not appear to be sustainable or in line with international law or with social science and management science standards (Lempert 2017).

A number of authors addressed these shortcomings by accentuating the need to align subnational indicator systems so they can be aggregated at the national level into country-level SDG progress reporting, which will

contribute to bottom-up accountability and reporting on SDGs. National statistical and other data collection agencies should play a critical role in fostering this alignment (Thrift and Bizikova 2016).

In its attempt to integrate the SDGs into its fragility analysis and framework, the OECD signals new approaches to fragility aimed at promoting fresh thinking and new discourse on fragility and how to better track needs, aid flows and progress in achieving the SDGs in fragile situations (Grotenhuis 2016; Michel 2018). Nevertheless, the new approaches to fragility involve welcoming the inclusion of security and peace within the new SDG agenda (Grotenhuis 2016).

Although the international community widely recognizes the need for a data revolution and the importance of supporting developing countries in significantly increasing the availability of high-quality, timely, reliable data, many fragile countries have poor-quality data, which undermines their capacity to make knowledge-based decisions and report accurately about their needs and performance (Michel 2018).

Processes that build upon existing awareness of context, build national ownership, foster transparency, secure support in the face of difficult choices and build consensus over trade-offs are crucial. Conversely, unsustainable practices that are having devastating environmental and social impacts on the local level can engender tensions and local discontent. For the 2030 Agenda to move forward in fragile states, decision-making must address not only technical, but also political obstacles to development (Carius 2018).

Palestine Explained

As mentioned previously, the OECD and many international organizations consider Palestine to be a fragile state. Palestine's fragility involves various elements, including the Israeli occupation of the West Bank and blockade of the Gaza Strip and the political divide between Gaza and the West Bank. These factors limit the national government's (Palestinian Authority) control over land, borders and natural resources. The limited control over unpredictable, ever-changing situations makes monitoring and tracking achievement of the SDGs under unstable, fragile conditions challenging. Monitoring the SDGs under such conditions is not easy and requires special skills and tools that need to be considered during the monitoring and evaluation process.

The dispute over 'Palestine', the area between the Jordan River and the Mediterranean Sea, has been one of the most complex, pressing, on-going

disputes since the beginning of the 20th century. This dispute extends beyond geography and politics to include the narrative and terminology. Therefore, a brief historical and geographical background is necessary to provide a context and backdrop for this chapter.

With a total land area of 26,323 km², historical Palestine lies on the western edge of the Asian continent and the eastern end of the Mediterranean Sea. Syria and Lebanon bound it to the north, the Gulf of Aqaba and the Egyptian Sinai Peninsula to the south, the Mediterranean Sea to the west and Jordan to the east.

After the first World War, the leading colonial powers of Britain and France controlled the League of Nations, which divided the territories of the collapsed Ottoman Empire. The territory of 'historic Palestine' was granted to Great Britain as a mandate. After the first Arab-Israeli war and the proclamation of the state of Israel on 15 May 1948, historic Palestine was divided into three parts: the new Jewish state, which occupied 78 per cent of the territory of Palestine; the West Bank of the Jordan River and East Jerusalem, which the Hashemite Kingdom of Jordan annexed; and a strip surrounding Gaza City close to the borders of Egypt that Egypt controlled (Ayyash 1981; Hajjar and Beinun 1988).

Israeli forces occupied the 22 per cent that remained under Arab sovereignty after 1948 in the Six Day War in June 1967. Although the Israelis prefer to use the biblical names of 'Judea and Samaria' to describe the southern and northern mountains of the West Bank, the international community, represented by the United Nations, refers to the West Bank and Gaza Strip as 'the Occupied Palestinian Territory-Palestine'.

On 13 September 1993, Israel and the Palestinian Liberation Organization (PLO) signed the Oslo I Accord, officially called the Declaration of Principles on Interim Self-Government Arrangements or Declaration of Principles, which was an attempt to establish a framework that would lead to resolution of the on-going Israeli-Palestinian conflict (BBC 2001). The agreement established the Palestinian Authority in May 1994, which has since taken on some civil and security responsibilities.

According to the agreement, the West Bank was divided into three zones: Areas A, B and C (PASSIA 2012) (figure 14.2). In Area A, which comprises 17.2 per cent of the West Bank and includes the major cities and villages, the Palestinian Authority has full security and civil responsibility, but Israel retains authority over movement into and out of these areas. In Area B, which comprises 23.8 per cent of the West Bank and includes most Palestinian villages, the Palestinian Authority has civil authority and responsibility for public order, and Israel maintains a security presence and

'overriding security responsibility'. In the remaining 59 per cent of the West Bank, Area C, Israel maintains security and civil powers.

After the Palestinian Authority took over responsibility for administration after the Oslo Accords, the Palestinian Territory was divided into 16 governorates (districts): 11 in the West Bank and five in the Gaza Strip.

The following sections explain the main drivers of the fragility that characterize Palestine.

Governance

The Israeli occupation of the West Bank and blockade of the Gaza Strip are the key drivers of fragility (Dittli 2011; World Bank 2012). The sectarian division between the Fatah movement led by Mahmoud Abbas, chairman of the PLO and president of the Palestinian Authority, and the Hamas movement exacerbates this fragility. Since June 2007, the Palestinian Authority, led by President Abbas, has governed the West Bank, and Hamas has ruled the Gaza Strip (UNDP 2010). The division between Hamas and Fatah has prevented elections from being held. The last presidential election was in January 2005, and the last legislative election was in January 2006, which Hamas won. The legislative council has not met since 2007, principally because of factional strife but also because Israel has imprisoned a number of its members (CEC 2020).

The political divide between Gaza and the West Bank is used as the pretext and justification for various forms of extra-legal activities or violations of human rights norms. 'It is highly important to note the bitter consequences of the [factional] divide [and]...the damage caused to the practice of human rights and freedoms, as a result of that divide' (ICHR 2012, 20).

As described before, the Palestinian Authority has jurisdiction over approximately 38 per cent of the West Bank territory, with the rest, including borders and water resources, under Israeli control. Lacking many of the

Figure 14.2 Map of Areas A, B and C in the West Bank



Source: PASSIA (2012).

instruments of a sovereign state, the Palestinian Authority is constrained in terms of what it can do to reduce and respond to fragility. In Hamas-run Gaza, the writ of the Palestinian Authority does not hold, and Israel controls the land and sea borders of the coastal strip except for a narrow land border with Egypt that is not always open, even for pedestrian traffic. In addition to these restrictions, Israel has imposed tight border controls and has limited access to coastal fishing areas and to farmland along its border (OCHA-OPT 2013).

Duplication of governance structures in the West Bank and in Gaza, expiry of presidential and legislative mandates and paralysis of the Palestinian Legislative Council all affect the legitimacy of the Palestinian Authority (Newton 2013), which the Palestinian Authority's limited ability to provide quality services for citizens in the West Bank and its absence from Gaza compound (World Bank 2011).

Economy

The Israeli occupation (and the imposed restrictions on access and movement), the fragmentation of economic space within the West Bank and between the West Bank and Gaza and the limited resource base due to Israeli control of 62 per cent of the territory of the West Bank are significant factors that hamper investment and undermine economic growth. Because of the Israeli restrictions, inefficiency and a lack of competitiveness characterize the Palestinian economy (World Bank 2012).

The Palestinian Authority's fiscal crisis also contributes to fragility. The inability of the Palestinian Authority to pay regular salaries to employees adds to the instability of more than 1 million Palestinians who depend directly and indirectly on government salaries (Portland Trust 2016). Moreover, the Palestinian Authority's chronic fiscal deficit has resulted in the accumulation of debt in the private sector and local banks (Flassbeck, Kaczmarczyk and Paetz 2018).

Environment

The environmental context in Palestine is extremely challenging. Some resources are severely degraded, access to others is limited or denied and certain ecosystems are on the brink of collapse. Population density and protracted conflict aggravate the situation, and the Middle East is highly vulnerable to climate change because of the risk of desertification and prolonged and recurring droughts (PNA 2012).

Under the Oslo Accords, most environmental problems are a shared responsibility between Israel and the Palestinian Authority. In practice, the Palestinian Authority's weak institutional capacity and Israel's control over water resources and most rural areas (which fall under Area C) hamper environmental management (World Bank 2009).

Dependence on Foreign Aid

The tendency of fragile states to depend on external donor aid has been growing since 2000 (OECD DAC 2012). Those states depend on official development assistance, which constitutes their largest source of finance, followed by remittances and foreign direct investment.

Palestine is considered one of the 'donor darlings' that receive half of all official development assistance directed to fragile states. With an official development assistance-to-gross domestic product ratio of 25.5 per cent, Palestine is ranked 12th among the world's most aid-dependent countries and economies (OECD 2018). This dependence has led the Palestinian economy to be structurally dependent on foreign aid (Devoir and Tartir 2009). Moreover, despite the massive amount of aid that Palestine receives, it is still unclear whether the existing coordination structure in Palestine is efficient in responding to local needs and priorities, not to mention fulfilling commitments to achieving the SDGs.

SDGs Efforts in Palestine

Having provided a detailed background of Palestine and the challenges it faces as a fragile state, we examine now the experience of Palestine in implementing and monitoring the SDGs.

Similar to other countries, and as part of its efforts to build institutions for statehood, the Palestinian government committed to working towards achieving the MDGs by aligning the priorities of government ministries and institutions to achievement of development in general, including fulfilment of the MDGs. Despite notable progress on various fronts and sectors, maintaining vulnerability to the Israeli policies that limit implementation of projects in Area C, East Jerusalem and the besieged Gaza Strip undermined achievement of the MDGs in Palestine. Israeli control over Palestinian land and resources has limited the capacity of the Palestinian Authority to advance further on the path towards building an independent Palestinian state and towards the development of Palestine that could benefit all regions and social groups (Palestinian Authority 2012).

The experience of the Palestinian government and other societal actors in implementing the MDGs was instrumental in shaping the Palestinian approach and efforts related to achievement and monitoring of the SDGs. A number of key lessons were recorded from the MDG experience⁴.

- *Inclusivity and partnership.* It was clear from implementation of the MDGs that the government cannot achieve its development agenda without real, inclusive partnership with key actors, including civil society, private sector, international community and most importantly citizens. For this reason, it was deemed critical to ensure that representatives of civil society and the private sector are on the national team to mainstream SDGs in their sectors in Palestine.
- Related to the previous point, *public awareness and buy-in* of the MDGs was limited. Actions undertaken were mostly ad hoc and focused on internal (governmental) awareness raising. Actions during the MDG period focused on building awareness and commitment among main stakeholders (government, civil society, private sector), with minimal attention to raising awareness of the general public, which limited buy-in and support of citizens for fulfilling the MDGs.
- *Alignment with local and national strategies.* Effective implementation of the global development agenda, whether MDGs or SDGs, entails harmonizing local and national priorities with the global agenda. The MDGs were adopted in parallel to, and often in isolation from, preparation of sectoral and national strategies, leading to limited coordination between sectors and interventions. To avoid this, the Palestinian government prioritized localization of the global agenda by including the SDGs in the National Policy Agenda (NPA) (2017–22).
- *Unified monitoring and evaluation system.* Lack of a national systematic monitoring and evaluation mechanism hindered tracking of and reporting on achievement of the MDGs. The absence of timely, reliable data during the period limited the ability of the government to report achievement of the MDGs.

⁴ Members of the Palestinian Authority National SDG Team shared these lessons with the author in interviews.

These lessons informed the Palestinian government's planning for and approach to implementing the SDGs and institutionalization of tracking, monitoring and reporting on the SDGs. A different, inclusive approach was followed after 2015 to lay the ground for implementation and monitoring of SDGs (PMO 2018; SAACB 2018).

Institutional Setup

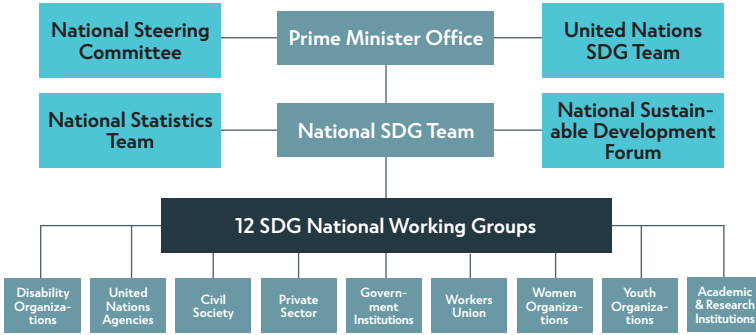
Since the time of the MDGs, Palestine has made enormous efforts to increase the capacity of national institutions and state structures and to mobilize local and international partnerships to support Palestine in achieving the SDGs in line with national development priorities.

Learning from the MDG experience, the government put off its highest priorities, first strengthening partnerships with relevant stakeholders by allowing as many stakeholders as possible to engage in follow-up and implementation of the SDGs.

The Palestinian Council of Ministers issued a decree on 19 February 2016, to form the National SDG Team to lead and coordinate the national effort to implement the SDGs under the leadership of the Prime Minister's Office, a reflection of interest and commitment at the highest official level. The National SDG Team was tasked with coordinating implementation and follow-up of the SDGs among all stakeholders, including from civil society, the private sector and the international community. As previously mentioned, establishment of the National SDG Team was based on lessons learned from the MDGs to encourage partners to participate in monitoring and implementing the SDGs at all stages of the process, from prioritization to implementation and evaluation (figure 14.3). The responsibility of the national team is to determine sustainable development priorities in Palestine and integrate them into the national framework for planning and budgeting processes, in addition to leading and coordinating preparation of national reviews of progress towards the SDGs.

Twelve SDG working groups were established to support the National SDG Team. The responsibility of the working groups, which relevant governmental institutions lead in close collaboration with relevant United Nations partner agencies, is to track the progress of their specific goal. The working groups are composed of representatives from civil society, the private sector and academic institutions. Each working group is tasked with one of the SDGs, with the exception of two working groups, one of which merged SDG 1 (No Poverty) and SDG 10 (Reduced Inequalities) and one of which merged the environment-related SDGs (12, 13, 14, 15). Because SDG 17

Figure 14.3 Institutional Arrangement for Follow-Up and Implementation of the SDGs



Source: PMO (2018).

(Partnerships) was seen to be a cross-cutting topic that concerns all groups, it was included in the mandate of all 12 groups. The National SDG Team designated focal points for each working group responsible for leading and coordinating national efforts to pursue and implement the SDGs. At the same time, the United Nations Resident Coordinator in Palestine identified focal points from respective United Nations agencies to support the SDG working groups (table 14.1). United Nations agencies were instrumental in supporting the government by assisting in follow-up, implementation and technical support to achievement of the SDGs.

The constitution of the National SDG Team and the associated SDG working groups offered an excellent opportunity to engage more than 300 stakeholders from various sectors in follow-up and implementation of the SDGs.

Ownership and Localization of the SDGs

As part of its efforts to achieve statehood, the Palestinian government has made serious efforts to integrate international and regional development agendas, including the SDGs, into national strategies and policies.

The launching of the 2030 Agenda coincided with preparation of the NPA 2017–22, which offered an opportunity to work towards incorporating the SDGs into the NPA from the onset of deliberations. At a sectoral level, ministries and government institutions have been asked to consider the SDGs in developing their sectoral and cross-sectoral strategies.

Table 14.1 Responsible Government Ministry and Supporting United Nations Counterpart Relevant to the Sustainable Development Goals

Goal	Lead Ministry	Counterpart Agency
1, 10	Ministry of Social Development	UNDP, UNICEF
2	Ministry of Agriculture	Food and Agriculture Organization of the United Nations
3	Ministry of Health	World Health Organization
4	Ministry of Education and Higher Education	United Nations Educational, Scientific and Cultural Organization
5	Ministry of Women's Affairs	UN Women
6	Palestinian Water Authority	UNICEF
7	Palestinian Energy and Natural Resources Authority	UNDP
8	Ministry of Labour, Ministry of the National Economy	International Labour Organization
9	Ministry of the National Economy, Ministry of Public Works and Housing	United Nations Office for Project Services
11	Ministry of Local Government, Ministry of Public Works and Housing	UN Habitat
12–15	Environmental Quality Agency	UNDP
16	Ministry of Justice	UNDP
17	Prime Minister Office	United Nations Special Coordinator for the Middle East Peace Process

Source: PMO (2018).

Note: UNDP = United Nations Development Programme; UNICEF = United Nations International Children's Emergency Fund.

NPA's Putting Citizens First is a national programme of action for Palestine focusing on the rights of citizens to freedom, justice, basic services, economic opportunities, safety and prosperity. A quick review of the NPA reveals that, in theory, the overall framework of the strategy is consistent with the transformational paradigm advocated in the 2030 Agenda, which is focused on marginalized and vulnerable groups and the 'leave no one behind' principle. The central focus of the NPA on citizens reflects the

government's priorities of strengthening its response to citizens and their needs while making greater efforts to raise the standard of living for the most marginalized despite the obstacles that the occupation poses. As stated in the Voluntary National Review (PMO 2018), realizing the principle of 'leave no one behind' in a country that lives under a protracted, ongoing occupation that imposes restrictions on its government and hinders its access to serve its people is a challenge.

After the broad consultation process to develop the NPA, it was apparent that, during the short term of the NPA (2017–22), Palestine will face difficulties in prioritizing all of the SDG targets because of the complex, fragile political and socioeconomic context and fiscal constraints. As a result, the NPA noted clearly that 'we must acknowledge that sustainable development cannot be achieved under Israeli occupation and without control over Area C's vast resources' (PMO 2016). It also emphasizes that sustainable development requires a holistic approach that cross-cuts a wide array of interventions, actions and priorities, all underpinned by the critical needs of the Palestinian population exacerbated by decades of Israeli military occupation.

With this understanding, and learning from the MDG experience, the Palestinian government embarked on localizing the SDGs in the Palestinian context in two steps: identifying national priorities from the SDGs and integrating priority targets into the NPA and sectoral strategies. The SDGs were therefore nationalized and integrated into national planning rather than adopted as a national agenda. Based on the comparison between the SDGs and the priorities of the NPA that the National SDG Team conducted, it was decided that, over the time frame of the NPA (2017–22), the State of Palestine would focus on 75 of the 169 interlinked targets that cover various domains and development sectors, although when considering the sectoral and cross-sectoral strategies, which focus on an additional set of targets, the total number of targets is 105. Priority targets were identified based on several factors, including the needs and priorities of the population, resource availability, the planning time frame and the specific conditions under which Palestinian institutions operate under the occupation, which limits their access to land and communities and their ability to provide services to the Palestinian people.

Data Availability and Monitoring of the SDGs

One of the key lessons learned from the MDG experience is the importance of having a national monitoring system to track and assess achievement

of the development agenda. For this purpose, the Palestinian government delegated the Palestinian Central Bureau of Statistics (PCBS) to localize and monitor the SDG indicators in cooperation with relevant ministries and national institutions and in coordination with international organizations, especially United Nations agencies concerned with these indicators (SAACB 2018).

It is highly challenging to collect the data needed to monitor SDGs in Palestine because of restrictions imposed on access of the Palestinian government to certain areas of Palestine (particularly East Jerusalem and Area C), lack of financial resources (e.g. conducting large-scale surveys, paying salaries of civil servants), and the limited capacity of some institutions and individuals.

To overcome the unavailability of data, the PCBS strengthened and institutionalized the administrative records system as a main source of data to bridge data gaps and to standardize data sources and methodologies. In addition, organizational and institutional changes were introduced to the PCBS by forming a working team on SDGs and instating an independent department for statistical control, central records and administrative data. The mandate of this department is to collect, compile, publish and document statistics from administrative records of various public and private institutions and combine these records with data from findings of surveys and censuses. In addition, the PCBS updated its data management structure to fulfil its vision of integrating data producers into the PCBS system and expanding the use of data from administrative records for statistical purposes. This made it easier to monitor SDG indicators and statistical monitoring indicators at the national level in the social, economic and environmental sectors according to the national strategic vision and goals.

To allow for measurement of SDGs, the PCBS and the National SDG Team have established a database of available indicators and source of data. One hundred nine of 244 (45 per cent) sustainable development indicators were found to be available, whereas some of the indicators are not categorized or stratified according to, for example, region, gender, or age (PMO 2018). Moreover, some of the SDG indicators were found to be irrelevant to the Palestinian context and hard to measure and achieve. For example, lack of Palestinian Authority control over water resources, coastal lines and most of its territory that is still under Israeli control limits achievement of SDG 14 (marine and coastal ecosystems) and SDG 15 (sustainable forest management). The challenges that the Palestinian Authority faces in building critical large-scale infrastructure such as airports, industrial parks and wastewater treatment because its financial resources and sovereignty

over land and natural resources are limited are making it difficult to achieve SDG 9 (industry, innovation and infrastructure)⁵, as well as multilateral environmental agreements on hazardous waste and other chemicals (SDG 12) and free trade and access to international market (SDG 17).

Producing the Voluntary National Review Report

One of the milestones that demonstrated the commitment of the Palestinian government to sustainable development was the National SDG Team's production of the Voluntary National Review Report in June 2018, which was presented to the international community at the United Nations High-Level Political Forum on Sustainable Development. The goal of the Voluntary National Review was to provide detailed information about the status of Palestine's progress towards implementing and achieving the SDGs, highlight the challenges that undermine pursuit of the SDGs, enhance stakeholder engagement in the follow-up and implementation of the SDGs and support creation of a developmental roadmap for implementation of the SDGs in Palestine.

The Prime Minister's Office led the review process, which ensured high-level participation and buy-in, as well as national ownership of the report's findings and outcomes. A steering committee of senior officials and the National SDG Team, which includes representatives from 24 governmental, non-governmental and private sector organizations, supported the Prime Minister's Office's leadership of the process. The 12 SDG working groups were tasked with mapping and monitoring the targets and indicators relevant to their associated SDGs, and the PCBS provided the needed statistical data (PMO 2018).

Concluding Remarks

The goal of this chapter was to highlight the experience and progress of Palestine as an example of a fragile state pursuing and monitoring the SDGs. It covered the unique challenges that hinder Palestine's ability to achieve sustainable development. At the top of these challenges is the ongoing colonization and occupation of Palestine and the deliberate policies and

⁵ External control over Palestine's borders and lack of an inland airport hinder the mobility of Palestinians. Two years after its construction, the Gaza airport was bombed, forcing Palestinians to travel to Amman, Jordan, to connect to the world.

restrictions of the Israeli occupation. Related to this challenge, some of the SDG targets cannot be achieved or monitored, including those related to conservation and sustainable use of the oceans, seas and marine resources for sustainable development; clean water and sanitation; the environment; trans-boundary issues; Area C; Bedouin communities and grazing land (PMO 2018).

Other obstacles to achieving the SDGs in Palestine include increasing pressure on basic services due to high natural population growth (~2.8 per cent), lack of resources to finance the development agenda due to the inability of the Palestinian people to access and control their natural resources and the sharp decline in the volume of foreign aid to Palestine in recent years, the need to strengthen national institutional capacity and coordination in following up on and implementing a sustainable development plan, and lack of data to measure progress towards achievement of many SDGs despite significant efforts made in this regard (as detailed in the previous sections).

Despite these obstacles, there is room, albeit limited, to advance the development agenda. The Palestine Voluntary National Review Report outlined follow-up mechanisms that could be useful for other countries that face similar conditions, including (PMO 2018):

- Focus on following up implementation of national sustainable development priorities to accelerate progress on implementation of the SDGs while strengthening coordination and joint work of stakeholders from different sectors to ensure efficiency of work and thus obtain the best results at the lowest costs.
- Promote community participation in follow-up and implementation of SDGs through development and application of a national strategy to strengthen partnerships with all parties and groups concerned with SDGs.
- Develop and implement a national mechanism for monitoring and reporting on implementation of SDGs and conduct an annual review in coordination and cooperation with all stakeholders.
- Cooperate and coordinate with all stakeholders to provide data on SDG indicators.
- Mobilize and develop partnerships at the regional and international levels on national sustainable development priorities.
- Raise institutional and community awareness of SDGs at the national and local levels.
- Promote participation of the media in following up implementation of SDGs.

Finally, the main underlying question that this chapter attempted to address is what and how much can be done to achieve sustainable development in a fragile state under a protracted occupation. Although I do not claim that this chapter provided complete answers to this question, I hope the case of Palestine can bring the attention of the academic and professional communities to the challenges that fragile countries face and the need for a more tailored, adapted evaluation and implementation approach that considers political, social and economic factors affecting the development agenda in these contexts.

References

- Ayyash, A. I. 1981. Israeli planning policy in the occupied territories. *Journal of Palestine Studies* 11 (1): 111–123.
- BBC (British Broadcasting Company). 2001. Text: 1993 Declaration of principles. Declaration of principles on interim self-government arrangements. London: BBC.
- Brinkerhoff, D.W. 2005. Rebuilding governance in failed states and post-conflict societies: Core concepts and cross-cutting themes. *Public Administration and Development* 25: 3–14.
- Brinkerhoff, D.W. 2010. Developing capacity in fragile states. *Public Administration and Development: The International Journal of Management Research and Practice* 30 (1): 66–78.
- Carius, A. I. 2018. *A foreign policy perspective on the SDGs*. Berlin: Adelphi.
- Carment, D.T. 2017. *The CIPF fragility index: New trends and categorizations: A 2017 country indicators for foreign policy report*. Ottawa: Norman Paterson School of International Affairs, Carleton University.
- CEC (Central Election Commission). 2020. Central Election Commission, Palestine. <http://www.elections.ps/tabid/815/language/en-US/Default.aspx>
- Chaitanya Kanuri, A. R. 2016. Getting started with the SDGs in cities: A guide for stakeholders. Sustainable Development Solutions Network. <https://resources.unsdsn.org/getting-started-with-the-sdgs-in-cities>
- Devoir, J., and A. Tartir. 2009. *Tracking external donor funding to Palestinian non-governmental organizations in the West Bank and Gaza 1999–2008*. Ramallah: Palestine Economic Policy Research Institute.
- Dittli, R. 2011. *Conflict sensitivity assessment: International assistance in Gaza: Aiding fragmentation or unity? A view from inside Gaza*. Basel, Switzerland: Center for Peacebuilding.
- Eizenstat, S., J. Porter and J. Weinstein. 2005. Rebuilding weak states. *Foreign Affairs* January/February: 134–146.
- Flassbeck, H., P. Kaczmarczyk and M. Paetz. 2018. *Macroeconomic structure, financial markets, and the financing of government activity: Lessons for Palestine*. Ramallah: Palestine Economic Policy Research Institute.

- François, M., and I. Sud. 2006. Promoting stability and development in fragile and failed states. *Development Policy Review* 24: 141–160.
- Greenhill, R. 2016. Where next for development effectiveness? Leaving no one behind. CAPE Conference.
- Grotenhuis, R. 2016. *Nation-building as necessary effort in fragile states*. Amsterdam: Amsterdam University Press.
- Hagesteijn, R. 2008. Early states and 'fragile states': Opportunities for conceptual synergy. *Social Evolution & History* 7: 1.
- Hajjar, Lisa, and Joel Beinin. 1988. Palestine for beginners. *Middle East Report* 154: 17–20.
- ICHR (Independent Commission for Human Rights). 2012. *The status of human rights in Palestine: Annual Report No. 18*, 18th ed., Ramallah: ICHR.
- Lempert, D. 2017. Testing the global community's Sustainable Development Goals (SDGs) against professional standards and international law. *Conscience* 18 (2): 111–175.
- Lister, S.W. 2005. Strengthening subnational administration in Afghanistan: Technical reform or state-building? *Public Administration and Development: The International Journal of Management Research and Practice* 25 (1): 39–48.
- Marcio Cruz, J. F. 2015. *Ending extreme poverty and sharing prosperity: Progress and policies*. Washington, DC: World Bank.
- Mcloughlin, C. 2010. *Topic guide on fragile states*. Birmingham, UK: Governance and Social Development Resource Centre.
- Michel, J. 2018. *Managing fragility and promoting resilience to advance peace, security, and sustainable development*. Washington, DC: Center for Strategic and International Studies.
- Miyazawa, I. 2012. *What are Sustainable Development Goals?* Kanagawa, Japan: Institute for Global Environmental Strategies.
- Newbrander, W.P.B. 2012. A tool for assessing management capacity at the decentralized level in a fragile state. *International Journal of Health Planning and Management* 4 (27): 276–294.
- Newton, C. 2013. Interview with UN Special Rapporteur Richard Falk: Legitimacy is the agency of history. *Truthout*. <https://tinyurl.com/wvj22m8t>
- OCHA-OPT (United Nations Office for the Coordination of Humanitarian Affairs Occupied Palestinian Territory). 2013. *Access restricted areas in the Gaza Strip*. Jerusalem: OCHA.
- OECD (Organisation for Economic Co-operation and Development). 2014. *Fragile states 2014: Domestic revenue mobilisation in fragile states*. Paris: OECD.
- OECD. 2018. *States of fragility 2018*. Paris: OECD Publishing.
- OECD DAC (Development Assistance Committee). 2012. *Fragile states 2013: Resource flows and trends in a shifting world*. Paris: OECD.
- Ottaway, M. 2002. Rebuilding state institutions in collapsed states. *Development and Change* 33 (5): 1001–1023.

- Palestinian Authority. 2012. *The national strategy to achieve the MDGs by 2015*. Ramallah: Palestinian Authority.
- PASSIA (Palestinian Academic Society for the Study of International Affairs). 2012. *Area C: The key to the two-state solution*. Jerusalem: PASSIA.
- Patton, M., and D. Horton. 2009. *Utilization-focused evaluation for agricultural innovation*. Rome: International Labor Accreditation Cooperation.
- PMO (Prime Minister's Office). 2016. *State of Palestine's 2017–22 National Policy Agenda: Putting Citizens First (final draft)*. Ramallah: PMO.
- PMO (Prime Minister's Office). 2018. *Palestinian national voluntary review on the implementation of the 2030 Agenda*. Ramallah: State of Palestine.
- PNA (Palestinian National Authority). 2012. *Sustainable development under Israeli occupation: Achievements and challenges*. Ramallah: PNA.
- Portland Trust. 2016. *Palestinian economic bulletin*. Ramallah: Portland Trust.
- Rakodi, C. 2001. Forget planning, put politics first? Priorities for urban management in developing countries. *International Journal of Applied Earth Observation and Geoinformation* 3 (3): 209–223.
- Roberts, A. 1990. Prolonged military occupation: The Israeli-occupied territories since 1967. *American Journal of International Law* 84 (1): 44–103.
- SAACB (State Audit and Administrative Control Bureau). 2018. *Review of the Palestinian government preparedness for the Sustainable Development Goals final report*. Ramallah: SAACB.
- Sekhar, C. S. C. 2010. Fragile states: The role of social, political, and economic factors. *Journal of Developing Societies* 26 (3): 263–293.
- Thrift, C., and L. Bizikova. 2016. *Bottom-up accountability for Sustainable Development Goals (SDGs) Are subnational indicator systems aligned?* Winnipeg, Canada: International Institute for Sustainable Development.
- UN DESA (United Nations Department for Economic and Social Affairs). 2017. *The Sustainable Development Goals report 2017*. New York: UN DESA.
- UNDP (United Nations Development Programme). 2010. *Human development report 2009/10: Occupied Palestinian territory : Investing in human security for a future state*. New York: UNDP.
- World Bank. 2009. *Assessment of restrictions on Palestinian water sector development*. Washington, DC: World Bank.
- World Bank. 2011. *West Bank and Gaza institutional capacity*. Washington, DC: World Bank.
- World Bank. 2012. *Towards economic sustainability of a future Palestinian state: Promoting private sector-led growth*. Washington, DC: World Bank.

PART V

APPROACHES
AND
METHODS





CHAPTER 15

Complex Systems, Development Trajectories and Theories of Change

AARON E. ZAZUETA, NIMA BAHRAMALIAN, THUY THU LE,
JOHANNES DOBINGER AND EKO RUDDY CAHYADI

Abstract. This chapter presents a case study of the application of Complex Adaptive Systems thinking to the planning, monitoring and evaluation of transformational interventions. The chapter presents a methodology to develop a robust understanding of the dynamics of the system targeted by a development intervention and to understand the ways in and extent to which a development intervention interacts to modify the development trajectory of such a system. It also describes the lessons learned in the ongoing learning process seeking to develop transformational theories of change. It draws on mixed methods that include different conceptual frameworks, analytical tools and information-gathering techniques. The approach we adopted is different from other systems thinking-inspired theories-of-change approaches in two ways. First, instead of focusing on the transformation of a system, our approach focuses on how to steer a system development trajectory that is consistent with a set of long-term objectives that are typically broadly articulated. The second important difference is that, unlike other systems-based theories of change, which often focus on transformation pathways that identify likely sequences of developmental stages (or conditions), our approach focuses on affecting the most influential conditions to steer the system development trajectory in the direction of the stated objectives. Our approach also focuses on monitoring the most effectual conditions to continually assess the extent and direction to which change takes place.

Introduction

This chapter presents the lessons learned in the ongoing effort to develop theories of change to aid transformational interventions that was first reported in Zazueta, Le and Bahramalian (2021) in the *American Journal of Evaluation* (the AJE article). Although this chapter and the AJE article both concern the United Nations Industrial Development Organization (UNIDO) technical assistance programme to the fisheries and aquaculture sector in Indonesia and are funded by the government of Switzerland, the AJE article refers to the evaluation of the first phase of the programme, called SMART-Fish, whereas this chapter addresses the development of a theory of change that is being used in the design, implementation and evaluation of the programme's second phase, which is delivered within the framework of UNIDO's Global Quality and Standards Programme (GQSP). The evaluation and the development of the theory of change both draw on complex adaptive systems (CAS) for their theoretical framework. They both adopted an approach based on a robust understanding of the dynamics of the system targeted by a development intervention and on the way the development intervention interacts with such systems (Garcia and Zazueta 2015). Like the AJE article, this chapter documents an experience that used mixed methods, but the methodological tools used in the process discussed in this chapter are more diverse and reliable:

- The analytical methods used are different. A key aspect of the shared overall approach is a focus on the interaction of the different conditions that contribute to change. In the evaluation discussed in the AJE article, we based our analysis on identification of the causal link among conditions leading to system change, which we did jointly with the project manager. In the development of the second phase, as discussed in this chapter, we asked stakeholders (some 30 persons) to define the causal links but also to weigh the influence of each condition. Although both processes were consultative, the process documented in this chapter was much more participative.
- For the AJE article, we used network analysis to identify the conditions' causal links in the system. For this chapter, we used the Decision-Making Trial and Evaluation Laboratory (DEMATEL), a mathematical tool that traces links across the system to the n link. When combined with the ponderation of each of the causal links, this made for a much powerful and reliable method.

- In the experience presented in this chapter, by using influence mapping, we were also able to systematically assess the role of each condition in the system, going from highly causal – and critical to redirect the system trajectory – to highly effectual – an indicator that change is taking place.
- In this chapter, we also used the Knowledge, Aspirations, Skills and Attitude model of the Bennett Hierarchy (Rockwell and Bennett 2004) and the Unified Theory of Acceptance and Use of Technology Utility (UTAUT) for planning interventions that promote behavioural changes that contribute to the identified highly influential conditions (Venkatesh, Thong and Xu 2016).

During the development process of the second phase, the lockdown caused by the COVID-19 pandemic required the project preparation team to use a variety of techniques to gather information, which included stakeholder workshops and web-based focus group discussions using the application Mural¹ and web-based questionnaires. The project design team also used visualization, using diagrams to facilitate analysis and communication, and consulted the relevant scientific and technical literature. This mix of conceptual frameworks, analytical and planning tools and data gathering and communication techniques was selected to match the objectives of the different steps in the development of a comprehensive model to steer the system development trajectory to the intended long-term objectives.

The approach was applied to the preparation of a project aimed at increasing market access of Indonesian fisheries and aquaculture products by improving compliance capacity with international quality standards. This project is taking place within the context of the GQSP, which is a strategic cooperative effort of the Swiss State Secretariat for Economic Affairs and UNIDO. The GQSP programme aims at promoting trade and competitiveness by strengthening quality and standards compliance of enterprises. The GQSP programme in Indonesia builds on a previous project implemented by UNIDO – SMART-Fish, which is also funded by the Swiss State Secretariat for Economic Affairs to strengthen the trade capacity of the seaweed, pangasius and pole and line tuna value chains. The GQSP aims at supporting aquaculture farms and fish processing firms, as well as intermediary actors such as collectors along the pangasius, shrimp, milkfish, catfish and seaweed value chains to meet their market requirements in terms of

¹ Mural is a digital workspace for visual collaboration. See <https://www.mural.co/>.

quality and compliance with technical regulation and standards. The project also supports enhancement of the overall quality infrastructure system of the aquaculture sector, which includes standards and regulation, capacity building of institutions supporting upgrading of and compliance with market requirements and the services to verify such compliance (conformity assessment bodies). The project works with various stakeholder groups to promote Culture for Quality, which refers to overall awareness of and demand for quality products by markets and demand for quality-related services by enterprises.

Development of the theory of change (ToC) for the second phase of the Indonesia project was a collaborative effort that included project managers and evaluators. The intention was to develop a model that could be used as a tool for adaptive management of the project during implementation and for project evaluation. Strengthening of the national quality system requires involvement of multiple ministries, academia, business associations and other civil society organizations, as well as consideration of several domains, including policy, technology, institutional, scientific and financial. Changes must take place at various levels: national, provincial and local. It is expected that changes will be phased in over a time frame that extends beyond the duration of the project.

One important aspect of the approach is incorporation of stakeholder and technical perspectives at each step in the process. The original intention of the authors was to apply the methodology in Indonesia through face-to-face stakeholder workshops, focus groups and interviews. This was not possible because of the COVID-19 pandemic. Instead, most of the work was conducted remotely using virtual tools to interact with participating stakeholders. It provided an opportunity to assess the impact of a pandemic on development processes. To accomplish this, the authors obtained information pertaining to conditions before and in the early stages of the COVID-19 outbreak when developing the baseline for the most influential conditions affecting the development trajectory.

The chapter is developed in three sections. The first pertains to identification of key CAS concepts used as tools to delimit the boundaries of the system that the project addresses and that consist of a set of assumptions on the causal links between different elements of the system, helping explain its dynamics. The second outlines the steps followed in developing the ToC. The third presents conclusions and lessons learned from the experience.

CAS Conceptual Tools

Particularly crucial in ToCs is clarity as to how the different elements of a system interact to generate the desired outcomes (Davies 2018). A ToC is a theory of why and how an initiative works (Weiss 1997). It explains how short-term interventions contribute to long-term objectives. Thus, ToCs are useful in designing and assessing the extent to which interventions are likely to contribute to long-term goals such as the United Nations Sustainable Development Goals (Selomane et al. 2019).

Useful ToCs start with a good understanding of the system that interventions target for change. System-wide changes typically require attention to complex processes. Among the many approaches to complex systems thinking, the CAS proponents have developed a set of concepts to understand and model the interlinked elements of a social system. CAS are dynamic systems that result from the behaviour of autonomous agents who are adapting to behaviours of other agents. The interactions between agents permeate the aspects of the system and are the source of unpredictability in the system (Dooley 1996). CAS are adaptive because they can reorganize their parts and learn when facing internal or external drivers (Anderies, Janssen and Ostrom 2004; Dooley 1996). The concepts that CAS scholars propose present a framework on which to build models consisting of propositions that can guide a learning process to help understand and modify the factors enabling or hampering the agents' behaviours. Particularly useful concepts to construct ToCs of system-wide changes include the concepts of system boundaries, domains, scales, agents, adaptive behaviour, emergence and system development trajectory.

The notion that CAS are composed of nested and interconnected subsystems helps define the boundaries of a system or a subsystem based on the intensity of links among the parts (Ostrom 2009). Subsystems that are interconnected through multiple links help identify the aspects of the phenomena relevant to the policy or long-term objectives. The domains pertain to the broad areas of concern relating to a phenomenon.

Domains are areas of knowledge (fields of cognition) or activity characterized by a set of concepts, terminology and behaviours (Couture and Valcartier 2007). The policy and regulatory, economic, institutional, financial, technological and sociocultural domains are frequently relevant in development initiatives (Zazueta 2017). Domains provide frameworks to help identify the critical conditions that can enable or hamper behavioural change consistent with given development trajectory. Examples of such conditions include the presence of policies and regulations that provide

incentives for behaviours consistent with the long-term policy goals, markets that recognize production processes and practices that contribute to the long-term policy goals and institutions capable of performing certain functions (UNIDO 2019). Domains are subsystems nested within the broader system that cut across scales; thus, identification of key domains pertaining to a given long-term objective is helpful in tracing the relevant system interactions and dynamics across different scales.

Scales represent levels across the system. Scales can be spatial or temporal and have quantitative and qualitative dimensions. For example, spatial scales go from local to national and global (in each case, there is an increase in the territory covered) and can pertain to different aspects of the phenomena, which could include administrative structures, ethnic boundaries, ecological boundaries and market systems (Gibson, Ostrom and Ahn 2000). Temporal scales can be short, medium or long term and can pertain to different phenomena such as frequency of occurrences across time and linearity or non-linearity in change. The links between scales make it possible for changes originating at one scale to trigger developments at other scales and across the system (Selomane et al. 2019).

The agents (or system components) and their behaviour underlie the phenomena that encompass the system. CAS scholars assume that systems operate through the actions and reactions of the agents (the agents' adaptive behaviour). Although agents command different resources, and the conditions in the various domains influence them differently, they are linked, directly or through other agents. The aggregated adaptive behaviour of the agents responding to each other and to other factors external to the system result in the emergence of system-level phenomena that are often different from the behaviours of the agents. The agents' adaptive behaviour can be in many forms, such as imitation, cooperation, conflict and coalitions, which feed back to influence other agents' reactions (Allen and Garmestani 2015; Holland 2006; Levin 2003).

The chain reaction of the adaptive behaviour of agents across scales and domains contributes to the complexity, non-linearity and unpredictability of the system (Holland 1995; 2006). Thus, in complex systems, outputs or results will not always correspond to inputs. Given these uncertainties, when dealing with complex systems, effective development interventions are those that mimic other agents in the system by adapting actions on the basis of information generated during the implementation process (Hartman and De Roo 2013). Adaptive management entails identification of long-term goals that can help guide the direction or trajectory of an intervention (Allen and Garmestani 2015). Adaptive management

also requires identification of conditions that are likely to enable human behaviours across the system that are consistent with the desired trajectory. This requires clear understanding of how the different conditions interact and influence the development trajectory of the system. Considering the uncertainty and unpredictability inherent in complex systems, this step will require articulation of a set of hypotheses and specific objectives that can be tracked and monitored regularly to make necessary adjustments based on information generated during implementation. To measure changes, it will also be necessary to identify the starting point for each condition and to develop indicators to trace changes and developments during implementation.

Steps in Developing the ToC

We followed eight steps in developing the ToC:

1. Formation of an expert group that included representatives from the different stakeholders of the project
2. Articulation of the long-term objectives, with engagement of the technical working group
3. Identification of conditions that would enable a development trajectory consistent with the long-term objectives
4. Analysis of the chains of causality within the system and identification of the most influential conditions across domains
5. Establishment of baseline of conditions
6. Identification of indicators for the most influential conditions and of a baseline
7. Development of strategies to contribute to the most influential conditions affecting the development trajectory of the system
8. Building of hypotheses for evaluation of results and designing an appropriate impact evaluation strategy

Step 1: Expert Group

In the first step, project management identified a group of experts and informed actors from the different sectors that could bring in different perspectives to enrich the exercise. These included members of industry associations, technical staff at the Ministry of Maritime Affairs and Fisheries, researchers and university lecturers and individual experts in aquaculture value chains that the project targeted.

This expert group, consisting of 40 persons, was formalized as the ToC Working Group and tasked to act as an advisory group to the Project Steering Committee, which consists of representatives of the Indonesian government, the Swiss State Secretariat for Economic Affairs (the donor) and UNIDO. The functions of the ToC Working Group included:

- Support of the project management team in designing the ToC and project interventions
- Periodic review of and updates to the ToC
- Periodic review of project results and providing advice for adaptive management
- Support of the project mid-term evaluation by revisiting the ToC and evaluating the project's contribution to changes in key conditions
- Support of the final evaluation of the project by assessing the project's contribution to changes in key and targeted conditions and helping identify lessons learned

The functions of the ToC Working Group are stipulated in terms of reference. To ensure continuity and meaningful engagement of the experts in the working group, the project provided a certificate of membership and recognized contributions of the Working Group in the project reports, but the most important incentive was the opportunity to contribute to and eventually influence policy in the sector in which they are all involved and feel passionate about. ToC Working Group participants are also becoming familiar with the concept of ToC as a tool for project planning, monitoring and evaluation and may be able to apply it in their work.

Step 2: Long-Term System-Wide Objectives

This step aimed at reviewing key objectives and goals of the sector as articulated in the policies and strategies of the government of Indonesia, as well as aspirations of the industry. Before the workshop, the project preparation team reviewed such policies and objectives to expedite the review process.

During discussions of the ToC Working Group, it became clear that, in the case of the aquaculture sector in Indonesia, the objectives were quite diverse, covering social, economic and environmental aspects, in line with the Sustainable Development Goal framework. Therefore, rather than agreeing on a specific objective, the group reached a consensus on the overall direction of the desired development. This was stipulated in the

three key words equity, wealth and sustainability, which also reflect key policy objectives of the government of Indonesia.

Step 3: Identifying Key Enabling Conditions and Domains of Change

The project preparation team presented a set of domains to the ToC Working Group as a starting point to help kick off a brainstorming session to identify key enabling conditions that would enable progress in the trajectory of the long-term goals identified in step 1. These domains were governance, knowledge and innovation, finance, production and markets.

This brainstorming was conducted in a face-to-face workshop in which ToC Working Group members were divided into subgroups and each group assigned two or three domains. Each group was then instructed to identify key conditions within the assigned domain that would steer behaviour in the trajectory identified in step 1. The subgroups were asked to limit each domain to five to six enabling conditions. Once breakout sessions were over, the whole group reconvened to review and validate key conditions that each subgroup identified. The review included removing duplicates, revising wording of conditions to ensure a consistent style of presenting conditions and adding new conditions if necessary; 27 enabling conditions were identified in the five domains (see annex for a detailed list of domains and conditions.)

The domain of production refers to a subsystem of activities that produce, transform and market aquaculture and seafood products. The expert groups identified six key conditions that constitute the positive contribution of this domain to the overall objectives of the sector. Unsurprisingly, awareness of sustainable development issues (P5) was selected as a key prerequisite for any progressive development towards sustainability, and any progressive change in the domain of production will ultimately depend on good practices, which will have a strong impact on sustainability of the sector if the actors along the value chain (input suppliers, farmers, processors and traders) adopt it at scale. Therefore, capacity to apply, for example, good aquaculture practices, good manufacturing practices, and good hygienic practices was defined as a key condition (P1). Moreover, capacity to meet market requirements and demand (P3), especially considering two other conditions in the domain of market (a growing and diversifying market for sustainable seafood (M5), market incentives along the value chain for quality, sustainability and equity (M4)), was identified as a key condition. Good production infrastructure (P4), suitable business

models (P2) and access to suitable production inputs (P6) were other key conditions identified in this domain.

The domain of markets was defined as the subsystem that facilitates transactions along the aquaculture value chain to meet demand for aquaculture products – be it sales of raw material for processing, of fresh fish into local markets or of processed seafood for exports. In such a subsystem, the key conditions were therefore appropriate trade infrastructure and logistics (M1); symmetric information on demand (price, quality, quantity; M2); suitable value chain coordination models (M3); market incentives along the value chain for quality, sustainability and equity (M4) and a growing and diversifying market for sustainable seafood (M5).

The expert team was sensitized and aware of system-wide changes needed at the sector level to achieve stated long-term objectives and the various trends and mega-trends, such as rural-urban migration and growth of digital technologies used in the sector. In light of this, openness to new knowledge and innovation (K1) was selected as a key condition in the domain of knowledge and innovation. The expert group agreed that a set of three conditions; good, strong research and development and training institutions (K2); cooperation between university and research and development institutions, government and industry (K3) and a good incentive system for innovation (K4) together create the needed knowledge and support the necessary changes at the system level, which transfers the new knowledge to relevant actors and encourages application of new knowledge and skills. The ability and responsiveness of the aquaculture sector to adapt to changes in circumstances, new developments and external shocks and megatrends (K5) and the availability of evidence to support policy and decision-making (K6) were considered two other important conditions. Such conditions were considered to be closely connected to the domain referred to as knowledge and innovation.

The expert group recognized that any changes in behaviour and any undertaking would require investment in new practices, processes, skills and technologies and suitable financing options. Therefore, a subsystem of finance was defined to include all the interconnected conditions referring to availability and accessibility of financial resources, feasibility of investments and capabilities associated with new investments. These included bankable proposals for investment in sustainable aquaculture (F5), appropriate financial business models along the value chain (F2) and financial literacy of different actors (F3). Moreover, the expert group decided that two conditions (presence of financial institutions promoting investments with impact on sustainability (F1) and good coordination among different

funding institutions (F4)) determine availability and accessibility of financial resources.

The expert group identified five conditions that referred to overall sector management capacity and governance towards a sustainable trajectory. Conducive regulation and standards for ease of doing sustainable business (G1) and a reward and punishment system for (un)sustainable practices (G4) underpin this capacity. Participatory approaches to governance and a strong and effective civil society (G2) was considered an important condition that enables interaction between the sector and the larger society. Recognizing the multiple streams of policy that affect the aquaculture sector, such as labour, agriculture, industry, trade and education, the expert group defined coherence and harmonized policies and capacity to implement policy (G5) and good coordination and cooperation among public and private institutions for implementation of policy (G3) as two other key conditions.

Step 4: Mapping Influence Between Conditions

4.1. Identifying Direct Influence Between Conditions

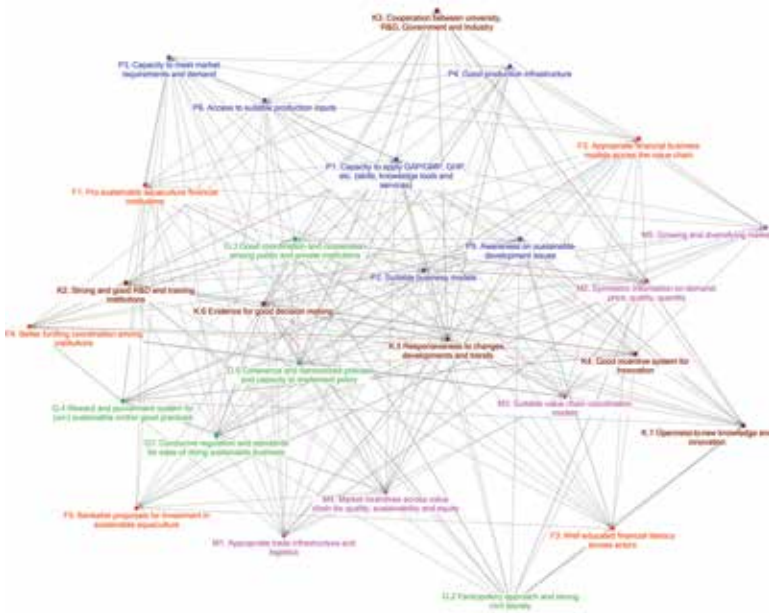
In the first stage, the whole group was engaged to identify direct influence links between conditions. To do so, the workshop facilitators went through every condition, domain by domain, and asked expert groups to identify the conditions that influence the selected condition the most², noting only direct influences between conditions (0 = no direct link, 1 = direct link); 247 direct links between the 27 conditions were identified. The result of substep 4.1 was subsequently diagrammed using NodeXL for network analysis and is presented in the network map in figure 15.1.

4.2. Evaluation of Strength of Direct Influence Between Conditions

Once the direct influences were mapped, the next substep was to evaluate the strength of each direct influence (0 = no influence (already identified in

² This question could have been posed the other way around as well, that is, picking a condition and asking the expert group to identify conditions that the selected condition influenced, but one can always assume some level of influence between all conditions in a system, and depending on the expert background, one may have exaggerated expectations of influence of a condition on the system. Therefore, it was deemed more accurate to ask about conditions that directly affect a selected condition.

Figure 15.1 Network of Key Conditions for Transforming Aquaculture Sector in Indonesia Towards Sustainability, Wealth and Equity



Source: 2020 data from the project team.

previous stage), 1 = weak, 2 = medium, 3 = strong, 4 = very strong). Substep 4.2 was conducted using an online survey. The definition of strength levels was considered subjective, meaning that each respondent could understand and use terms such as ‘weak’, ‘medium’, ‘strong’ and ‘very strong’ differently. Therefore, the facilitation team explained to the ToC Working Group that, in this substep, it was particularly important for each respondent to apply their own definition consistently to have consistent responses throughout the survey and not to worry about consistency of definitions across respondents. Twenty-four experts completed the survey, which took approximately three hours. The facilitation team was available and could reach out to respondents to address enquiries on the questions. In addition, a WhatsApp group was formed to share questions and answers with Working Group members. Subsequently, a two-hour webinar was organized to review the results of first phase and to explain the next steps, including the survey. The result of this substep was 24 matrices of influence between conditions of the ToC.

4.3. Identifying the Most Influential Conditions

In a complex system in which conditions influence each other, directly and indirectly (through other conditions), understanding the most influential conditions is crucial for programmes and projects that intend to influence the trajectory of the system towards the long-term objectives. Because the project does not have the resources to develop detailed strategies for the 24 enabling conditions, identifying the most influential conditions will allow the design team to pay more attention to development of robust strategies to pursue the conditions that are most likely to steer the system towards the long-term objectives. The most influential conditions (or factors, as referred to in operations research literature) in a complex system were identified using the DEMATEL technique, a structural modelling approach that translates the interdependency relationships between conditions of a complex system into cause-and-effect groups. As such, it determines whether a condition is a driver or cause of change or a result or effect of other conditions. In addition, DEMATEL identifies the most important conditions of a complex system with the help of an impact relation diagram by calculating the total routes (direct and indirect) through which a condition influences other conditions and the system as a whole (Shafiee, Lofti and Saleh 2014).

The project preparation team used classic DEMATEL³ to analyse the strength of the causal links resulting from the survey responses obtained in substep 4.2. The individual direct influence matrices of substep 4.2 were collected into an aggregate direct influence matrix, and then a total influence matrix was elaborated by summing all direct and indirect influences between conditions.

The application of the DEMATEL method is summarized as follows.

Generate the individual direct-influence matrix (X). The following calculations used the data in the non-negative matrix attained from respondents after assessing the relationships between n conditions. The matrix captures the responses of m respondents who were asked to indicate the direct influence of condition C_i on condition C_j using an integer score of 0, 1, 2, 3 or 4, representing no influence, low influence, medium influence, high influence and very high influence, respectively.

³ The project used the process that Si et al. (2018) described for a classic DEMATEL technique.

x_{ij} indicates the degree to which the respondent believes condition i affects condition j . For $i = j$, the diagonal elements are set to zero. For each respondent, an n -by- n non-negative matrix can be established as:

$$X^k = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nn} \end{bmatrix},$$

where k is the number of respondents with $1 \leq k \leq m$, and n is the number of conditions. Thus, X_1, X_2, \dots, X_m are individual direct influence matrices from m respondents. Twenty-four experts ($m = 24$) assessed the strength of the relationship between 27 identified conditions ($n = 27$). In other words, 24 individual direct influence 27×27 matrices were developed in this stage.

Computation of group direct-influence matrix. To incorporate all opinions from all respondents, the average matrix is computed as:

$$x_{ij} = \frac{\sum_{k=1}^m x_{ij}^k}{m}$$

Computation of normalized direct influence matrix. Normalize direct-influence matrix D as follows:

$$D = X.S$$

Where

$$S = \frac{1}{\max_{1 \leq i \leq n} \sum_{j=1}^n x_{ij}}$$

Computation of total-influence matrix. Total relation matrix (T) is defined as $T = D(I - D)^{-1}$, where I is the identity matrix.

Calculation of 'Prominence' and 'Relation' values. The total-influence matrix shows all the direct and indirect influences from each condition on all other conditions in the system. This will give us the essential elements to assess the importance and role of conditions in the system, referred to, respectively, as prominence and relation.

We first define R and C by $n \times 1$ and $1 \times n$ vectors representing the sum of rows and sum of columns, respectively, of the total relation matrix T . In other words, for each of the 27 conditions, R is the sum of all direct and

indirect effects that a condition has on other conditions in the system (sum of rows of total influence matrix). Similarly, C is the sum of all direct and indirect effects on a condition of all other conditions in the system (sum of columns of the total relation matrix) (table 15.1).

'Prominence' shows the strength of influences that a condition gives and receives (out-degree plus in-degree). Prominence is calculated for each condition by adding its R and C values. The higher the $(R + C)$ of a condition, the more central a role the condition plays in the system.

'Relation' shows the net effect that the condition contributes to the system, which can be cause or effect. Relation is calculated for each condition by deducting C from R . Specifically, if $(R - C)$ of a condition is positive, it is a cause, whereas if $(R - C)$ is negative, a condition is an effect.

Influential relation map. In this substep, an influential relation map plot was illustrated by mapping conditions based on their prominence $(R + C)$ and relation $(R - C)$ values on a scatter plot (figure 15.2).

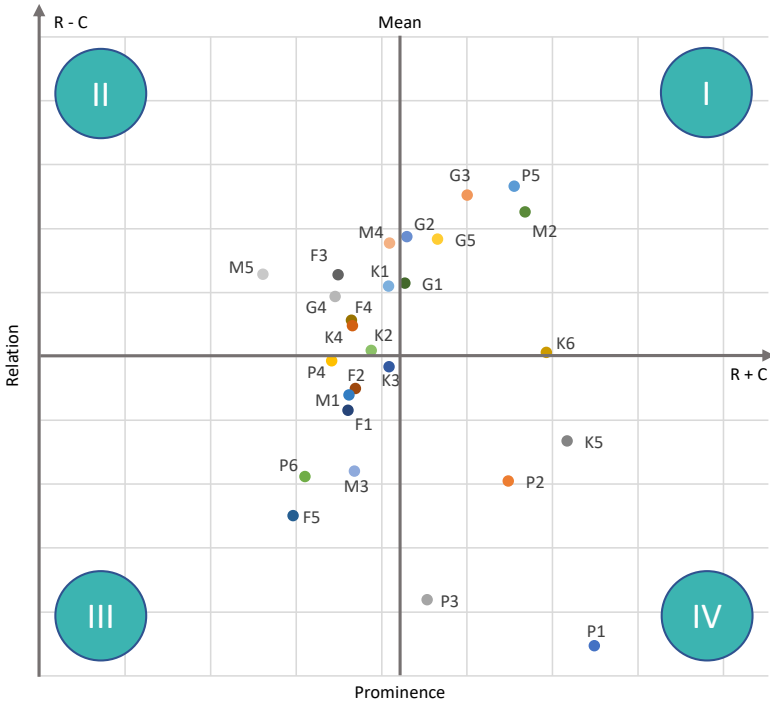
Si et al. (2018) suggest the following classification of the conditions in a complex system according to their quadrant location in the influential relation map diagram:

- Conditions in quadrant I are the most important conditions because they have high prominence and relation. For the purpose of the intervention design, these are conditions that the project should target to have the greatest influence on the trajectory of the system.
- Conditions in quadrant II are identified as autonomous driving conditions because they have low prominence but high relation. These conditions have a strong causal effect in the model. They are less connected in the system but have strong influence on the conditions they directly influence.
- Conditions in quadrant III are independent conditions that are relatively disconnected from the system because they have low prominence and low and negative relation. Although these conditions are relatively disconnected, they are relevant because they are considered necessary to achieve the long-term goal.
- Conditions in quadrant IV have high prominence but low and negative relation. These are referred to as impact factors. Other conditions strongly influence them. These conditions are indicators of the extent to which long-term impact is being achieved.

Table 15.1 Total Influence Matrix

	P1	P2	P3	P4	P5	P6	F1	F2	F3	F4	F5	G1	G2	G3	G4	G5	K1	K2	K3	K4	K5	K6	M1	M2	M3	M4	M5	R
P1	0.02	0.07	0.09	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.07	0.07	0.01	0.01	0.01	0.01	0.00	0.49
P2	0.10	0.02	0.09	0.01	0.01	0.02	0.02	0.08	0.01	0.07	0.09	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.07	0.08	0.01	0.00	0.00	0.88
P3	0.01	0.01	0.01	0.00	0.00	0.01	0.07	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18
P4	0.10	0.08	0.10	0.01	0.00	0.07	0.02	0.02	0.00	0.01	0.08	0.01	0.00	0.00	0.01	0.06	0.00	0.01	0.02	0.01	0.09	0.02	0.01	0.01	0.02	0.01	0.06	0.84
P5	0.17	0.12	0.14	0.03	0.03	0.05	0.05	0.09	0.07	0.08	0.05	0.09	0.03	0.08	0.08	0.02	0.09	0.10	0.10	0.08	0.13	0.11	0.04	0.03	0.10	0.08	0.01	2.05
P6	0.07	0.01	0.08	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.01	0.01	0.00	0.01	0.00	0.00	0.31
F1	0.10	0.08	0.10	0.07	0.00	0.08	0.01	0.01	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.01	0.07	0.01	0.07	0.00	0.00	0.69
F2	0.04	0.09	0.09	0.07	0.01	0.07	0.02	0.01	0.01	0.00	0.01	0.09	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.07	0.01	0.02	0.01	0.07	0.06	0.00	0.80
F3	0.11	0.10	0.05	0.02	0.07	0.02	0.09	0.09	0.01	0.02	0.10	0.02	0.06	0.01	0.01	0.01	0.07	0.02	0.02	0.02	0.09	0.08	0.03	0.02	0.03	0.02	0.00	1.19
F4	0.11	0.03	0.04	0.01	0.01	0.03	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.07	0.06	0.01	0.02	0.07	0.08	0.07	0.09	0.03	0.08	0.02	0.09	0.01	0.01	1.05
F5	0.01	0.01	0.01	0.00	0.00	0.01	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.12
G1	0.13	0.11	0.12	0.07	0.02	0.09	0.09	0.03	0.01	0.01	0.04	0.01	0.07	0.01	0.06	0.02	0.01	0.07	0.02	0.02	0.05	0.09	0.08	0.02	0.03	0.01	0.07	1.36
G2	0.08	0.11	0.06	0.03	0.08	0.03	0.04	0.08	0.03	0.04	0.04	0.04	0.03	0.08	0.02	0.09	0.08	0.09	0.04	0.07	0.11	0.11	0.03	0.08	0.05	0.02	0.01	1.54
G3	0.15	0.11	0.13	0.03	0.02	0.10	0.03	0.09	0.07	0.09	0.05	0.09	0.08	0.02	0.02	0.09	0.02	0.08	0.09	0.03	0.07	0.11	0.09	0.08	0.10	0.02	0.01	1.88
G4	0.12	0.10	0.11	0.01	0.07	0.02	0.08	0.02	0.01	0.01	0.03	0.01	0.01	0.01	0.02	0.01	0.07	0.02	0.02	0.07	0.04	0.03	0.02	0.03	0.07	0.06	0.10	1.10
G5	0.14	0.05	0.12	0.08	0.02	0.09	0.03	0.07	0.01	0.08	0.04	0.08	0.01	0.01	0.09	0.02	0.02	0.03	0.09	0.09	0.11	0.09	0.03	0.02	0.09	0.08	0.02	1.62
K1	0.13	0.05	0.05	0.02	0.08	0.02	0.07	0.03	0.07	0.02	0.03	0.03	0.07	0.02	0.01	0.02	0.02	0.08	0.03	0.02	0.11	0.10	0.03	0.08	0.08	0.01	0.00	1.30
K2	0.11	0.09	0.10	0.06	0.01	0.07	0.02	0.02	0.06	0.02	0.03	0.02	0.01	0.05	0.00	0.01	0.01	0.01	0.02	0.01	0.10	0.09	0.02	0.01	0.02	0.01	0.01	0.99
K3	0.12	0.03	0.04	0.02	0.02	0.07	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.08	0.08	0.02	0.02	0.02	0.10	0.10	0.02	0.07	0.03	0.01	0.00	0.98
K4	0.12	0.09	0.04	0.01	0.02	0.07	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.07	0.01	0.08	0.08	0.08	0.01	0.04	0.09	0.02	0.02	0.07	0.01	0.01	1.04
K5	0.12	0.10	0.11	0.07	0.00	0.03	0.08	0.02	0.01	0.07	0.10	0.09	0.06	0.01	0.01	0.01	0.01	0.07	0.07	0.07	0.04	0.03	0.08	0.02	0.09	0.00	0.01	1.21
K6	0.13	0.05	0.06	0.03	0.07	0.04	0.03	0.08	0.01	0.02	0.10	0.09	0.06	0.02	0.03	0.08	0.02	0.03	0.09	0.08	0.11	0.04	0.07	0.02	0.04	0.07	0.01	1.50
M1	0.09	0.02	0.10	0.01	0.01	0.08	0.01	0.06	0.01	0.01	0.02	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.07	0.01	0.01	0.07	0.08	0.01	0.00	0.75
M2	0.16	0.13	0.07	0.08	0.08	0.05	0.04	0.11	0.08	0.03	0.11	0.09	0.08	0.08	0.02	0.09	0.08	0.03	0.09	0.03	0.07	0.12	0.09	0.04	0.11	0.03	0.01	1.99
M3	0.08	0.01	0.02	0.01	0.01	0.06	0.00	0.00	0.01	0.01	0.06	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.07	0.01	0.00	0.00	0.47
M4	0.13	0.11	0.12	0.07	0.07	0.03	0.09	0.03	0.01	0.02	0.09	0.02	0.01	0.07	0.08	0.01	0.02	0.07	0.03	0.08	0.10	0.03	0.08	0.02	0.04	0.01	0.01	1.47
M5	0.11	0.09	0.04	0.02	0.01	0.02	0.08	0.08	0.00	0.01	0.09	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.08	0.07	0.08	0.07	0.02	0.01	0.03	0.01	0.00	0.97
C	2.75	1.86	2.08	0.87	0.72	1.24	1.11	1.05	0.55	0.77	1.36	0.78	0.60	0.62	0.63	0.70	0.74	0.95	1.06	0.79	1.87	1.46	1.05	0.85	1.37	0.58	0.33	

Source: Calculations by author based on 2020 survey data from the project team.

Figure 15.2 Influential Relation Map

Source: 2020 survey data from the project team.

The ToC Working Group concurred with the results of the analysis and identified the conditions in each of the four quadrants (table 15.2). Various conditions were identified in quadrant I to design project intervention strategies. P5⁴ (awareness of sustainable development issues), M2 (symmetric information on demand: price, quality and quantity) and G3 (good coordination and cooperation among public and private institutions) were identified as the most important factors driving change in the system because of their high prominence and high and positive relation. Conditions in this group also include G5 (coherence and harmonized policies and capacity to implement policy), G2 (participatory approach and strong civil society), G1 (conducive regulation and standards for ease of doing sustainable business) and K6 (evidence for good decision-making). These conditions

⁴ Working Group members were convinced of the importance of sustainability awareness (P5) and the need to promote knowledge in this regard.

Table 15.2 Classification of Conditions

Q-II Autonomous driving conditions	Q-I Most important conditions
<p>M4 Market incentives across value chain for quality, sustainability and equity</p> <p>K1 Openness to new knowledge and innovation</p> <p>F3 Well-educated financial literacy across actors</p> <p>M5 Growing and diversifying market</p> <p>G4 Reward and punishment system for (un)sustainable and/or good practices</p> <p>F4 Better funding coordination among institutions</p> <p>K4 Good incentive system for Innovation</p> <p>K2 Good, strong research and development and training institutions</p>	<p>P5 Awareness of sustainable development issues</p> <p>M2 Symmetric information on demand: price, quality, quantity</p> <p>G3 Good coordination and cooperation among public and private institutions</p> <p>G5 Coherence and harmonized policies and capacity to implement policy</p> <p>G2 Participatory approach and strong civil society</p> <p>G1 Conducive regulation and standards for ease of doing sustainable business</p> <p>K6 Evidence for good decision-making</p>
Q-III Independent conditions	Q-IV Impact conditions
<p>K3 Cooperation between academia, research and development, government and Industry</p> <p>F2 Appropriate financial business models along the value chain</p> <p>P4 Good production infrastructure</p> <p>M1 Appropriate trade infrastructure and logistics</p> <p>F1 Pro-sustainable aquaculture financial institutions</p> <p>M3 Suitable value chain coordination models</p> <p>P6 Access to suitable production inputs</p> <p>F5 Bankable proposals for investment in sustainable aquaculture</p>	<p>P1 Capacity to apply good practices</p> <p>P2 Suitable business models</p> <p>P3 Capacity to meet market requirements and demand</p> <p>K5 Responsiveness to changes, developments and trends</p>

Source: 2020 survey data by project team.

were identified as the key conditions driving the aquaculture sector towards competitiveness, equity and sustainability.

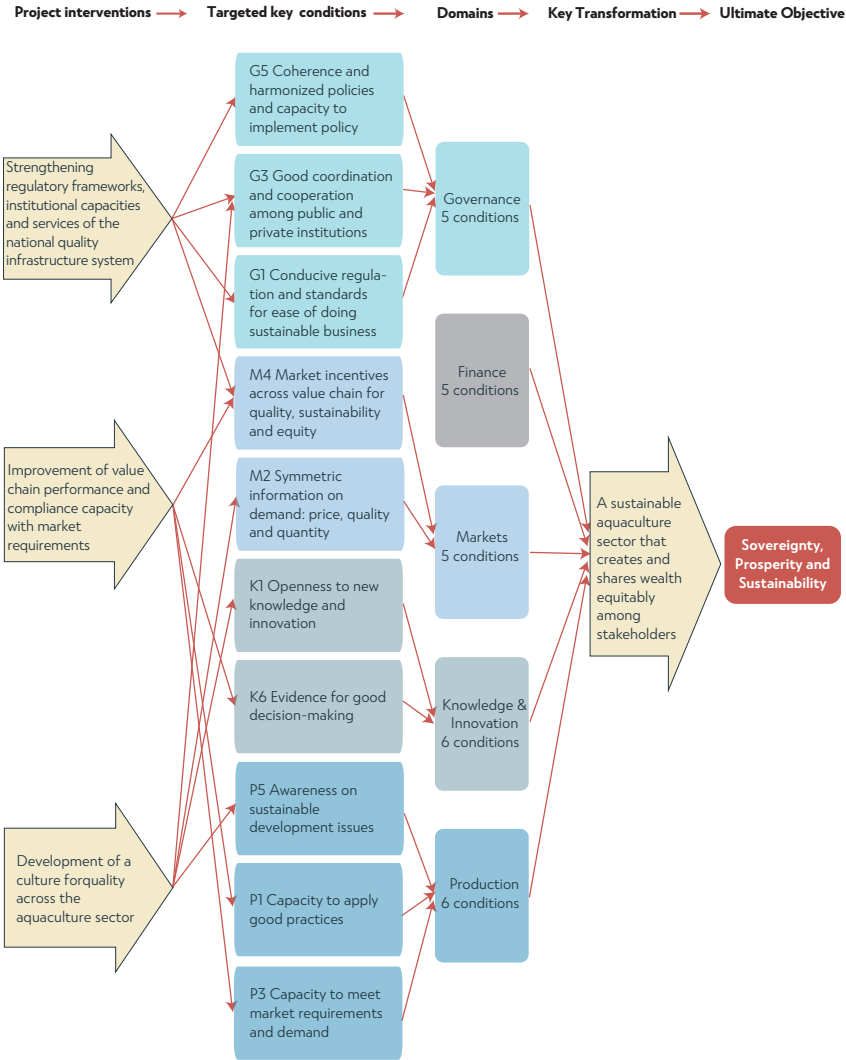
The ToC Working Group also selected conditions belonging to quadrant II that have a positive relation ($R - C$) but less than average prominence ($R + C$). These include M4 (market incentives across value chain for quality, sustainability, and equity), which had high positive relation, meaning that it is an important cause factor driving change in the sector. Similarly, K1 (openness to new knowledge and innovation), F3 (well-educated financial literacy across actors) and M5 (growing and diversifying market) were cause factors that played less prominent roles in the system.

Conditions in quadrant IV are prominent and important conditions for achieving the long-term objectives. Although other conditions influence them strongly, they have less impact on the other conditions in the system. For example, condition P1 (capacity to apply good practices) is where real change takes place when industry actors start applying business practices that, if adopted at scale, will achieve long-term objectives. The project therefore selected this condition too, to develop a targeted strategy of interventions that would lead to adoption of good practices along value chains (see step 7 for further details).

Step 4 was critical in this process because it identified 10 of the 27 initially identified key conditions on which the project could concentrate its resources to support the desired development trajectory. This does not mean that other conditions were not important, because they are also necessary to steer the system in the direction of the desired long-term objectives. It does mean that, if the causality assumptions in step 4 are correct, the 10 conditions could play a role in driving change in the other conditions, particularly conditions in quadrants II and IV.

Figure 15.3 illustrates the project's ToC and maps its interventions against the most important conditions identified in step 4. At the extreme right of figure 15.3 are the government of Indonesia's guiding principles of property, sovereignty and sustainability, which help guide the national development trajectory. The project seeks to support this development trajectory by contributing to the conditions that are conducive to a sustainable aquaculture sector that creates and shares wealth equitably. Although the ToC formulation process identified 27 conditions there, the project is focusing on 10 (in five domains) that were identified as the most influential conditions across the system (quadrant I) or that can be used as indicators of system-wide change (quadrant IV). At the extreme right are the project interventions, which include three sets of activities related to regulatory changes and capacity development, improvement of value chain

Figure 15.3 Global Quality and Standards Programme Theory of Change



Source: Authors.

performance and improvement of a culture for quality across aquaculture. Although figure 15.3 indicates a unidirectional chain of causality, the interplay of interactions among conditions is much more complex, with causality moving in different relations, as indicated in figure 15.2. Two key assumptions of the ToC are that the 10 selected conditions have a high level of influence and representation of change across the 27 conditions in the system and that progress in these 10 conditions is highly likely to steer the system in the desired development trajectory.

Step 5: Baseline Assessment

In this step, the ToC Working Group assessed the current status of each of the 27 conditions in the aquaculture sector in Indonesia with regard to the long-term objectives of creating wealth, equitable distribution of value across the industry and environmental sustainability. Information was collected using an online survey. Participants were asked to rate the state of each condition by selecting 0 = very bad, 1 = bad, 2 = medium, 3 = good or 4 = very good. The result of this assessment is the baseline against which changes in the state of conditions and the overall changes of the system can be measured.

Respondents were also asked to rate the state of each condition before and after the COVID-19 outbreak⁵. The recent global pandemic has affected aquaculture in Indonesia, as well as the global markets on which it depends. The question was intended to:

- better understand the impact of COVID-19 on the sector and key conditions
- test the robustness of the developed ToC
- potentially identify key areas of immediate policy intervention to counter negative impacts of the pandemic

Table 15.3 presents average expert scores for each condition before and after the COVID-19 pandemic outbreak and indicates whether each condition's score is below or above the average.

As indicated in table 15.3 and illustrated in figure 15.4, most conditions (20) that the ToC Working Group assessed have significantly worsened during the COVID-19 pandemic. Only in six conditions did the ToC Working

⁵ The survey was conducted in June 2020, only three months into the pandemic; therefore, the full short-term effects may not have been apparent and reflected in responses.

Table 15.3 Expert Group Assessment of Baselines

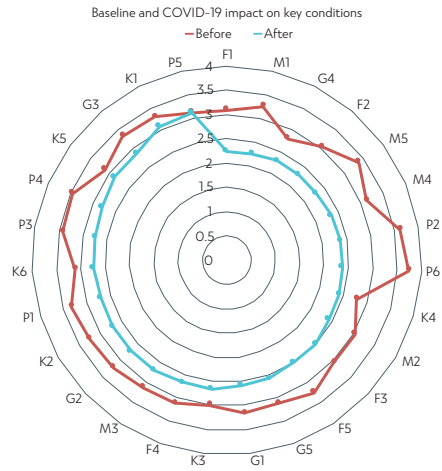
Quad-rant	Condi-tion	Description	Before COVID-19			After COVID-19		
			Score	Rank	State	Score	Rank	State
II	G4	Reward and punishment system for (un)sustainable and/or good practices	2.79	26	Below	2.29	23	Below
II	K4	Good incentive system for Innovation	2.79	26	Below	2.42	18	Below
III	K3	Cooperation between academia, research and development, government and Industry	3.00	25	Below	2.67	12	Above
III	F2	Appropriate financial business models along the value chain	3.04	22	Below	2.29	23	Below
I	M2	Symmetric information on demand; price, quality, quantity	3.04	22	Below	2.42	18	Below
II	F3	Well-educated financial literacy across actors	3.04	22	Below	2.50	16	Below
III	F1	Pro-sustainable aquaculture financial institutions	3.08	21	Below	2.25	26	Below
II	M4	Market incentives along value chain for quality, sustainability and equity	3.13	14	Below	2.33	22	Below
I	G5	Coherence and harmonized policies and capacity to implement policy	3.13	14	Below	2.58	14	Below
II	F4	Better funding coordination among institutions	3.13	14	Below	2.67	12	Above
III	M3	Suitable value chain coordination models	3.13	14	Below	2.71	8	Above
IV	K5	Evidence of good decision-making	3.13	14	Below	2.75	6	Above
IV	K5	Responsiveness to changes, developments and trends	3.13	14	Below	2.88	3	Above
I	P5	Awareness of sustainable development issues	3.13	14	Below	3.13	1	Above
I	G1	Conductive regulation and standards for ease of doing sustainable business	3.17	13	Below	2.58	14	Below
III	M1	Participatory approach and strong civil society	3.21	12	Above	2.71	8	Above
III	M1	Appropriate trade infrastructure and logistics	3.25	9	Above	2.25	26	Below
III	F5	Bankable proposals for investment in sustainable aquaculture	3.25	9	Above	2.50	16	Below
II	K2	Good, strong research and development and training institutions	3.25	9	Above	2.71	8	Above
II	K1	Openness to new knowledge and innovation	3.29	8	Above	3.08	2	Above
IV	P1	Capacity to apply good agricultural/manufacturing/hygienic/etc. practices, skills, knowledge tools and services)	3.33	6	Above	2.71	8	Above
I	G3	Good coordination and cooperation among public and private institutions	3.33	6	Above	2.88	3	Above
II	M5	Growing and diversifying market	3.38	5	Above	2.29	23	Below
IV	P3	Capacity to meet market requirements and demand	3.42	4	Above	2.75	6	Above
III	P4	Good production infrastructure	3.46	3	Above	2.79	5	Above
IV	P2	Suitable business models	3.63	2	Above	2.38	20	Below
III	P6	Access to suitable production inputs	3.75	1	Above	2.38	20	Below
		Average	3.20			2.59		

Source: 2020 survey data by project team.

Group assessment show a small difference (P5, K3, K4, K5, K6, M3). Four conditions had larger declines than others (a mean difference of one or more between the before and after COVID-19 assessments). The most vulnerable conditions and those that the pandemic has most affected include:

1. P6 (adequate access to production inputs). It is understandable that, during the pandemic, the government applied full and partial lockdown policies in some districts, including regions that supply inputs. Condition P6 fell in quadrant III, the quadrant in which conditions are more autonomous than the other conditions affecting the system. Nevertheless, this condition remains important because, without production inputs, the production process is halted, and the value chain is interrupted.
2. P2 (suitable business model). During the pandemic, the business environment changed and disrupted existing business models. Shocks to the supply chain were observed in supply and demand.
3. M5 (growing and diversified market). The pandemic led to a decline in global demand.
4. M1 (adequate trade infrastructure and logistics). During the pandemic and under the lockdown policies, transportation activities on some roads and at some airports and ports were strictly limited. This directly affected logistics services, on which export of seafood relies heavily.

Figure 15.4 State of Conditions Before and After COVID-19



Source: 2020 survey data by project team.

Step 6: Indicators for Monitoring Changes at System Level

The expert assessment of the baseline by the ToC Working Group should be examined on the basis of evidence from other sources. It is highly likely that, for some of the conditions, suitable and standard quantitative indicators are available and accessible. For example, condition M5 refers to a

growing and diversifying market that can be measured using national and international trade data, but for most of the conditions, such indicators are not readily available, and their production would require significant effort. To this end, three possibilities can be pursued in parallel:

- Consulting relevant literature to identify suitable indicators related to each condition
- Focused discussion with key stakeholders related to the specific condition
- Substantiation of the baseline scoring by providing examples and reasons

Step 7: Building Strategy

The next step was to develop specific strategies to target each of the 10 conditions identified as having the most influence in the system development trajectory. Whereas transformation at the aquaculture sector level was considered a complex adaptive process, changes in individual conditions were assumed to follow a less-complex process. Small groups (five to six persons) that had the relevant expertise on topics related to each of the 10 conditions completed this stage. The online tool Mural was used to facilitate the process.

The strategy-building process included three substeps: identifying influence pathways that drive the targeted condition, stakeholder analysis relevant to the condition and constructing a hypothesis for monitoring and evaluation of outcomes of the strategy to influence the condition. We provide one example of such a strategy-building process for improving condition P1: capacity to apply good aquaculture practices, good manufacturing practices, and good hygienic practices (skills, knowledge tools and services). This strategy involves UNIDO as an external agent and other national public and private stakeholders.

In general, a standard design process was followed, including:

7.1. Mapping Influence Pathways

Using the influential relation matrix (table 15.1), the project preparation team mapped the direct and indirect influencers separately from the other 26 conditions that were mapped. The top seven influencing conditions on P1 are listed in table 15.4. These conditions were considered when analysing the specific problems affecting the capacity of value chain operators to adopt good practices.

Table 15.4 Most Influential Conditions on P1 (Capacity to Apply Good Practices)

P5	Awareness of sustainable development issues
M2	Symmetric information on demand: price, quality, quantity
G3	Good coordination and cooperation among public and private institutions
G5	Coherence and harmonized policies and capacity to implement policy
M4	Market incentives along value chain for quality, sustainability and equity
K6	Evidence of good decision-making
G1	Conducive regulation and standards for ease of doing sustainable business

Source: Project preparation team (2020).

7.2. Stakeholder Analysis

Changes in each condition are causally linked to the behaviour of key stakeholders. The small group of experts conducted a stakeholder analysis for each of the conditions by asking participants to identify stakeholders at various levels of the administrative scale (city, province, country) and to identify motivators of each stakeholder. To facilitate the process, five pre-defined categories were used to identify the stakeholders: research and education, industry and private sector, government, non-governmental organizations, and sector organizations and professional bodies.

7.3. Strategy for Change

The process of achieving desired behavioural change was mapped based on the Knowledge, Aspirations, Skills and Attitude model of the Bennett Hierarchy (Rockwell and Bennett 2004). In this substep, the small group of experts was asked about the current activities and role of each stakeholder in promoting good aquaculture practices. In other words, the expert group identified activities of key stakeholders aimed at reaching out to farmers (discover), providing information and raising awareness of farmers (inform), building capacity of farmers and improving their skills (training) and supporting farmers in implementing good aquaculture practices and applying newly obtained skills (implementation).

Once stakeholder activities were mapped against the behavioural change process, gaps were identified that key stakeholders did not target systematically. The small working group was then asked to brainstorm on possible responses to the identified gaps, and a list of interventions was created

that each stakeholder could undertake and that could eventually contribute to improvement of the overall state of the key condition in question.

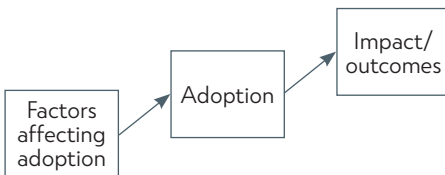
Step 8: Developing Hypothesis to Measure Outcomes and Impact

8.1. Conceptual Framework

The example that is described in step 7 refers to adoption of good practices by value chain actors (e.g. farmers). Therefore, the conceptual framework to measure success of the strategy requires identification of factors that affect adoption of good practices, the extent of adoption of good aquaculture practices by value chain actors and the extent to which good aquaculture practices contribute to desired outcomes consistent with long-term objectives in the aquaculture sector. The long-term objectives in question are related to creation of value and equitable distribution of wealth along the value chain while minimizing the environmental impact of farming activities.

This conceptual framework is depicted in figure 15.5.

Figure 15.5 Conceptual Framework for Impact Evaluation



Source: Authors.

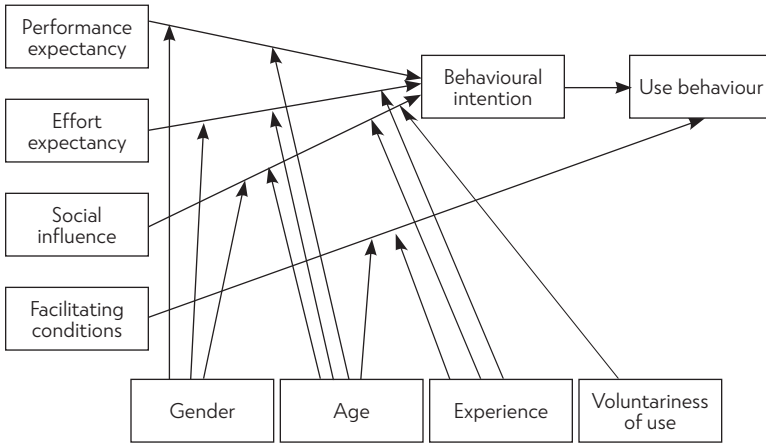
Adoption of good practices is considered a dynamic process and is influenced by multiple factors. The UTAUT utility (figure 15.6) provides a model to explain behavioural intention towards adoption of a technology⁶. The UTAUT provides a framework to analyse the factors that lead to user adoption of a practice by focusing on four constructs:

performance expectancy (PE), effort expectancy (EE), social influence (SI) and facilitating conditions (FC) (Williams, Rana and Dwivedi 2015).

Using UTAUT, the project establishes two hypotheses:

- The knowledge transfer strategy that key stakeholders develop and implement will result in farmers adopting good aquaculture practices.
- Adoption of good aquaculture practices contributes to desired impacts: improvement of farmer's livelihood, environmental sustainability and more balanced distribution of value created.

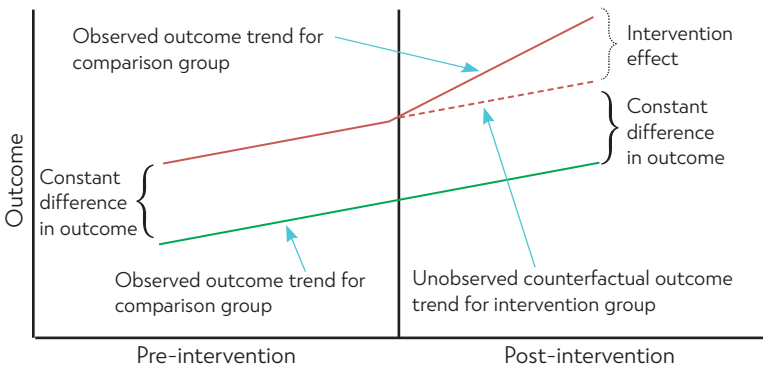
⁶ Technology in the general meaning of the word, including practices.

Figure 15.6 Unified Theory of Acceptance and Use of Technology

Source: Venkatesh et al. (2003).

8.2. Design of Evaluation Studies

Impact evaluation of the strategy is approached by comparing before and after project implementation. The evaluation design will apply a difference-in-differences (DID) model, which is used to estimate a causal effect using longitudinal data from treatment and control groups to obtain an appropriate counterfactual (figure 15.7). Three assumptions need to be

Figure 15.7 Difference-in-Difference Estimation, Graphical Explanation

Source: Columbia University Mailman School of Public Health (2019) (<https://tinyurl.com/ydm8t7v5>).

held: exchangeability, positivity and Stable Unit Treatment Value Assumption. The model also requires that:

- The intervention be unrelated to outcome at baseline (outcome did not determine allocation of intervention)
- Treatment or intervention and control group have parallel trends in outcome
- Composition of intervention and comparison groups be stable for repeated cross-sectional design
- There be no spill-over effects

DID will be implemented as an interaction term between time and treatment group dummy variables in a regression model:

$$Y = \beta_0 + \beta_1[\text{Time}] + \beta_2[\text{Intervention}] + \beta_3[\text{Time} \times \text{Intervention}] + \beta_4[\text{Covariates}] + \varepsilon$$

Where:

Y: Outcome

β_0 : Baseline average

β_1 : Time trend in control group

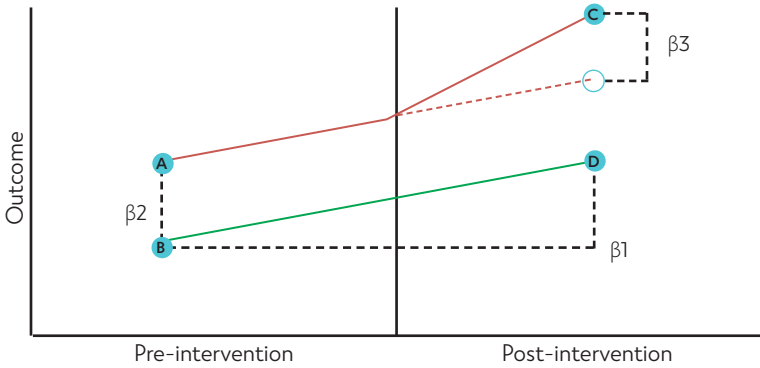
β_2 : Pre-intervention difference between two groups

β_3 : Differences in changes over time

The model can be explained through a graphical explanation (figure 15.8).

The benefit of adopting good practices will be assessed at two levels: group and individual. It is important to differentiate the impact of good practices of groups from that of individuals because the good practice may work only at the aggregate level and not at the individual, which adoption rates of individual farmers can affect. Therefore, assessing individual and group levels can reduce bias or overestimated impact. The benefit of the good practice is assumed to consist of production, quality, financial, social and environment aspects.

Data will be collected through a survey using structured questions and in focus group discussions. The survey will be conducted to collect quantitative data, whereas focus groups are required to capture qualitative opinions of farmers at the group level.

Figure 15.8 Model for Evaluating Benefits of Adoption of Good Practices

Source: Columbia University Mailman School of Public Health (2019) (<https://tinyurl.com/ydm8t7v5>)

Conclusions

Transformational development interventions confront major challenges stemming from highly unpredictable social systems due to the interaction of multiple and confounding factors and the emergence of conditions that are difficult to predict. Structural change is also likely to take place in time scales that go beyond the duration of the intervention. In this chapter, we presented an approach to help model the interactions among the key components of the system and to identify the conditions that are most likely to drive change towards a given development trajectory. We combined the use of CAS conceptual tools, network analysis and DEMATEL to model the interactions of the components of the system and to identify the most promising inflection points to steer the system development trajectory towards the desired long-term objective. This combined approach reduced drawbacks from using a single tool. CAS encourage deep observation of the system by identifying domains and help identify the conditions likely to contribute to the desired changes. Thus, CAS provided a comprehensive view of the system and of the extent of its complexity, although CAS are of limited utility for planning because it is not realistic that a project can directly target the multiple conditions and interactions of conditions across the system. Network analysis helped delineate the extent to which conditions influence each other. DEMATEL, by quantifying the effect of each condition, allowed us to identify cause factors, effect factors and the

most prominent factor. Combining CAS, network analysis and DEMATEL helped develop a ToC composed of a set of robust hypotheses of the most influential conditions affecting the system development trajectory. The approach followed also used a variety of techniques to gather information, ensuring that the perspectives of the key stakeholders and the appropriate technical knowledge were incorporated at every stage of the process. Stakeholder participation and sound interdisciplinary scientific and technical knowledge were particularly important when developing strategies to intervene in the most influential conditions.

The approach we presented combines CAS-inspired ToCs with the use of qualitative and quantitative tools to map and track change system inflection points that can help steer a system's development trajectories towards long-term transformational objectives, also allowing a more meaningful and robust evaluation than if only theory-based or quantitative approaches were used. The approach we have presented is different from other ToC approaches inspired by systems thinking in two ways. First, instead of focusing on the transformation of a system, our approach focuses on *how to steer a system development trajectory that is consistent with a set of long-term objectives that are typically broadly articulated*. The second important difference is that, unlike other systems-based ToCs, which often focus on transformation pathways that identify likely sequences of developmental stages (or conditions), our approach focuses on affecting the *most influential conditions* to steer the system development trajectory in the direction of the stated objectives.

We have developed the ToC; mapped the interactions across conditions and identified causal factors, effect factors and most prominent factors as part of the inception phase of project implementation. At the time of writing this chapter, the project team was in the process of using the UTAUT model to design specific strategies to affect most influential conditions. The project implementation team will develop indicators for each of the most influential conditions in the system to assess progress made in redirecting the development trajectory at the system level. The project implementation team is also developing UTAUT model indicators to help assess the extent of stakeholder behavioural change that is consistent with the desired development trajectory and to provide timely information for adaptive management. Whenever feasible, the project implementation team is using SMART-C (Specific, Measurable, Accountable, Realistic, Time bound and Challenging) key performance indicators to measure change in the key system conditions. Key performance indicators combined with the recurrent annual ToC Working Group reviews will function as an

early-warning system and learning mechanism to assist in adaptive management of the project during implementation.

On the practical side, the process of constructing a comprehensive ToC (which identifies conditions to steer the system in the desired trajectory, links between the conditions and the most influential conditions) is likely to take two to three months. A robust ToC requires in-depth understanding of the target system and its dynamics. To this end, it is necessary to build a common understanding among stakeholders through consultation, in-depth studies and review and discussion of the technical literature, but the project identification stage is often short and limited in resources, which does not allow for a full range of activities leading to development of a comprehensive model. In this example, the ToC was developed during the inception stage of the implementation phase. This can be a practical formula for similar projects, whereby the process can start in the project identification stage and continue to the inception stage. That said, establishing a ToC at any point in the life cycle of a project is useful, mainly because periodic revisiting of the ToC is necessary, especially considering the need for adaptive management. Although establishing a ToC at the outset should result in better strategic choices, doing so during the implementation phase can help project activities achieving better results.

The approach presented provides a framework particularly well suited to designing, managing and evaluating transformational interventions because it is based on a holistic understanding of the system an intervention seeks to influence. Other approaches seek to simplify complexity or design projects around specific development pathways. The model presented here helps identify a manageable number of conditions for intervention by embracing complexity by mapping the extent to which the key conditions interact and influence one another. We present an approach to provoke cause-and-effect cascades across the system that have a high likelihood of redirecting the system trajectory. The structure of the process is similar to that of a neural network, in which inputs provoke multiple interactions among system components, which are difficult to trace but result in a discernible set of outcomes (Shi 2019). A further comparison with neural networks is warranted. The approach we present engages technical and non-technical stakeholders from inception and throughout the project or programme cycle (including planning, monitoring and evaluation); this approach is likely to increase effectiveness and contribute to sustainability of the new trajectory by building stakeholder ownership of technically sound strategies and outcomes. Also, when incorporating multiple agents of development cooperation and sectoral ministries, the process helps

improve communication between agencies and identifies opportunities for coordination and collaboration. The ToC also provides a framework for periodic review of project accomplishments and development of performance and impact indicators to inform the adaptive management of the process.

References

- Allen, C.R., and A.S. Garmestani. 2015. Adaptive management of social-ecological systems: The path forward. In: *Adaptive management of social-ecological systems*, (ed.) A.S. Garmestani, 255–262. Springer Netherlands. https://doi.org/10.1007/978-94-017-9682-8_14
- Anderies, J., M. Janssen and E. Ostrom. 2004. A framework to analyze the robustness of social-ecological systems from an institutional perspective. *Ecology and Society* 9 (1). <https://doi.org/10.5751/ES-00610-090118>
- Couture, M., and D. Valcartier. 2007. *Complexity and chaos-state-of-the-art: Overview of theoretical concepts*. Ministry of National Defence, Canada.
- Davies, R. 2018. Representing theories of change: Technical challenges with evaluation consequences. *Journal of Development Effectiveness* 10 (4): 438–461. <https://doi.org/10.1080/19439342.2018.1526202>
- Dooley, K.J. 1996. A nominal definition of complex adaptive systems. *Chaos Network* 8 (1). <https://tinyurl.com/1lcoqhha>
- Garcia, J.R., and A. Zazueta. 2015. Going beyond mixed methods to mixed approaches: A systems perspective for asking the right questions. *IDS Bulletin* 46 (1): 30–43.
- Gibson, C.C., E. Ostrom and T.K. Ahn. 2000. The concept of scale and the human dimensions of global change: A survey. *Ecological Economics* 32 (2): 217–239.
- Hartman, S., and G. De Roo. 2013. Towards managing nonlinear regional development trajectories. *Environment and Planning C: Government and Policy* 31 (3): 556–570. <https://doi.org/10.1068%2Fc11203r>
- Holland, J. 1995. *Hidden order: How adaptation builds complexity*. Reading, MA: Addison Wesley Longman Publishing Co. Inc.
- Holland, J. 2006. Studying complex adaptive systems. *Journal of Systems Science and Complexity* 19: 1–8.
- Levin, S. 2003. Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bulletin of the American Mathematical Society* 40 (1): 3–19.
- Rockwell, K., and C. Bennett. 2004. *Targeting outcomes of programs: A hierarchy for targeting outcomes and evaluating their achievement*. Lincoln, NE: University of Nebraska. Faculty Publications, Agricultural Leadership, Education and Communication Department.
- Selomane, O., B. Reyers, R. Biggs et al. 2019. Harnessing Insights from social-ecological systems research for monitoring sustainable development. *Sustainability* 11 (4): 1190. <https://doi.org/10.3390/su11041190>

- Shafiee, M., F.H. Lofti and H. Saleh. 2014. Supply chain performance evaluation with data envelopment analysis and balanced scorecard approach. *Applied Mathematical Modelling* 38 (21–22): 5092–5112.
- Shi, Hualing. 2019. Application of back propagation (BP) neural network in marine regional economic forecast. *Journal of Coastal Research* 98 (S1): 67–70.
- Si, S.-L., X.Y. You, H.C. Liu et al. 2018. DEMATEL technique: A systematic review of the state-of-the-art literature on methodologies and applications. *Mathematical Problems in Engineering* 2018: 1–33. <https://doi.org/10.1155/2018/3696457>
- UNIDO (United Nations Industrial Development Organization). 2019. *Independent terminal evaluation Indonesia SMART-Fish programme. Increasing trade capacities of selected value chains within the fisheries sector in Indonesia*. Vienna: UNIDO.
- Venkatesh, V., M. G. Morris, G. B. Davis et al. 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly* 27 (3): 425–478. <https://doi.org/10.2307/30036540>
- Venkatesh, V., J.Y. Thong and X. Xu. 2016. Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems* 17 (5): 328–376.
- Weiss, C. H. 1997. How can theory-based evaluation make greater headway? *Evaluation Review* 21 (4): 501–524.
- Williams, M. D., N. P. Rana and Y. K. Dwivedi. 2015. The unified theory of acceptance and use of technology (UTAUT): A literature review. *Journal of Enterprise Information Management* 28 (3): 443–88.
- Zazueta, A. E. 2017. *Principles for the development of integrated climate change and chemicals and waste*. A STAP Information Document. Global Environment Facility Scientific and Technical Advisory Panel.
- Zazueta, A. E., T. T. Le and N. Bahramalian. 2021. Development trajectories and complex systems-informed theories of change. *American Journal of Evaluation* 42 (1): 110–129. <https://doi.org/10.1177/1098214020947782>

Annex: Domains and Conditions

Production

- P1. Capacity to apply good aquaculture practices, good manufacturing practices and good hygienic practices (skills, knowledge tools and services)
- P2. Suitable business models
- P3. Capacity to meet market requirements and demand
- P4: Good production infrastructure
- P5: Awareness of sustainable development issues
- P6: Access to suitable production inputs

Market

- M1. Appropriate trade infrastructure and logistics
- M2. Symmetric information on demand: price, quality, quantity
- M3: Suitable value chain coordination models
- M4: Market incentives along value chain for quality, sustainability and equity
- M5: Growing and diversifying market for sustainable seafood

Finance

- F1. Pro-sustainable aquaculture financial institutions
- F2. Appropriate financial business models along the value chain
- F3: Well-educated financial literacy across actors
- F4: Good coordination among funding institutions
- F5: Bankable proposals for investment in sustainable aquaculture

Governance

- G1. Conducive regulation and standards for ease of doing sustainable business
- G2. Participatory approach and strong civil society
- G3: Good coordination and cooperation among public and private institutions
- G4: Reward and punishment system for (un)sustainable and good practices
- G5: Coherence and harmonized policies and capacity to implement policy

Knowledge and Innovation

- K1. Openness to new knowledge and innovation
- K2: Good, strong research and development and training institutions
- K3: Cooperation between academia, research and development, government and industry
- K4: Good incentive system for Innovation
- K5: Responsiveness to changes, developments and trends
- K6: Evidence for good decision-making



CHAPTER 16

A Complexity-Based Meta-Theory of Change for Transformation Towards Green Energy

JONATHAN A. MORELL (JONNY)

Abstract. This chapter draws from complexity science to present a meta-theory of transformation that can be applied to discrete theories of change constructed to guide model building, methodology and data interpretation for evaluation of change efforts. The focus is on six specific behaviours of complex systems – stigmergy, attractors, emergence, phase transition, self-organization and path dependence. These can be invoked singly or in combination to understand pattern, predictability and how change happens. The importance of both ‘explanation’ and ‘prediction’ is woven into the discussion. A definition of ‘transformation’ is offered in which a qualitatively new reality becomes the default choice that constitutes a new normal. Indicators of transformation include measurable ranges (as opposed to specific values) for level of energy use and the time over which the change endures. Because complex systems behave as they do, the recommended theory of change is sparse; it has few well-defined elements or relationships among those elements. There is already good progress in the application of complexity to the evaluation of transformation. An argument is made that these efforts should be strengthened by deliberately incorporating what is known about complex system behaviour, and that, by so doing, both prediction and explanation would better serve the purpose of practical decision-making.

Why a Complexity-Based Meta-Theory of Transformation?

What follows is a theory about commonalities among theories of change, irrespective of their specific content. Put differently, I will articulate a theory about theories of change (a meta-theory). The meta-theory I will present will draw heavily on complexity science and will focus on transformation to a green energy future. I do not use the term 'complexity science' lightly. There is a deep epistemological literature concerning why the study of complexity deserves to be called a science (Phelan 2001).

The discussion will present many notions about action, measurement and causality that you may find uncomfortable, at odds with common sense or both (Morell 2017). I hope to convince you that, despite the discomfort and the challenges to common sense, the meta-theory I am about to present is worth taking seriously.

Any theory underlying evaluation of transformation to green energy must be judged with respect to its predictive power, explanatory power¹ and value as a useful guide to practical action.

We evaluate for instrumental and conceptual purposes. Success at both requires thought and action based on theories that respect the complex nature of change². To show why this is correct, I will proceed through three broad topics:

- concepts from complexity science that are relevant to evaluation theories of transformation
- the characteristics of 'transformation'
- how the previous two topics combine to form a meta-theory of transformation that can be applied to context-specific evaluations of transformation

Theories of action explain *how*. Theories of change explain *why* (Tyrrel 2019). This can be thought of as the difference between science and technology (Morell 1979). Technology turns to science when it is no longer able

¹ 'Explanation' and 'prediction' are not simple and obvious. For rigorous treatment, see Niiiniluoto (2019) and Shmueli (2011).

² The argument made here is part of a larger literature that draws on complexity to drive social science theory and methodology. For instance, Marion (1999) reinterprets many well-known organizational theories using a chaos and complexity framework.

to achieve desired results (when the implicit theory of ‘why’ on which a technology is based is no longer powerful or correct enough). This is exactly the problem that those seeking to effect transformation face. Their theories of action do not respect the complex nature of change.

Complex Behaviour, Not Complex Systems³

Evaluators must make operational decisions. How should programme theory be represented? What form should a logic model take? What methodology should be employed? What data should be collected? How should the data be analysed? How should the data be interpreted? They also need to make fuzzier, but nonetheless critical decisions: How to communicate to funders and other interested parties about realistic expectations for programme outcomes. How to help people understand the causal dynamics that drive programmes. How to explain the boundaries of what can and cannot be known about a programme’s consequences. Complexity-inspired answers to questions like these reside in knowing how complex systems *behave*, not what complex systems *are*.

The field of complexity is vast (Castellani 2009; 2014). It would be no more appropriate to say that ‘complexity’ is relevant to the evaluation of transformation than it would be to say that ‘statistics’ are relevant to the evaluation of transformation. What matters is: Which aspects of complexity are useful under which circumstances? There is no single answer to this question. This article is based on mine⁴.

Three themes that cut across the complexity landscape are pattern, predictability and how change happens. The rows of table 16.1 list the complex behaviours that I believe are most useful for evaluating transformation to green energy. The columns remind us that each of these complex behaviours may have implications for understanding some combination of the complexity themes: pattern, predictability and how change happens.

³ Three sources are particularly useful as entry points into the domain of complexity: New England Complex Systems Institute 2020, Santa Fe Institute 2020a, Systems Innovation 2020a.

⁴ For another classification that is well worth considering, see Boehnert (2020).

Table 16.1 Cross Reference: Complexity Themes and Complex Behaviours Useful in Evaluation

Complex behaviour	Theme in complexity science		
	Pattern	Predictability	How change happens
Stigmergy			
Attractors			
Emergence			
Phase transition			
Self-organization			
Sensitive dependence			

Relevant Complex Behaviours and Their Evaluation Implications

How do complex behaviours provide insights into pattern, predictability and how change happens? What are the implications for understanding transformation? To answer these questions, I will start by providing intuitive explanations for each row in table 16.1. After each explanation, I will discuss the evaluation implications of each complex behaviour. I will end by showing how complex behaviours come together to help understand pattern, predictability and how change happens.

Stigmergy

Stigmergy is a concept that was first developed to understand insect behaviour (Theraulaz and Bonabeau 1999) but has since been generalized to many human-scale situations in which changes in an environment serve as cues to direct the behaviour of subsequent actors (Parunak 2006). In a stigmergic process, even though there is no direct interaction with previous actors and no overall plan that any actor follows, a goal directed-type pattern is manifest. This happens because the 'plan' is embedded in the history of activity that independent actors encounter.

It may be an error to assume that a goal-directed theory of change must include deliberate planning. An alternative approach would be to consider whether the context is one in which independent actors react in specific ways to their environment, resulting in activity that looks as if it

were centrally coordinated. A stigmergic theory of change is particularly relevant to long timeline social changes that require multiple activities performed by multiple actors. This is because deliberate coordination among these actors is neither practical nor desirable (Morell 2018).

Attractors

Attractors are complexity's way of identifying 'where systems like to be', which is a loose anthropomorphic term, but one that provides an intuitive and accessible definition. More technically,

In the mathematical field of dynamical systems, an attractor is a set of numerical values towards which a system tends to evolve, for a wide variety of starting conditions of the system. System values that get close enough to the attractor values remain close even if slightly perturbed (Systems Innovation 2020b).

Social attractors define a specific subset of states that a social system may take, which corresponds to its normal behaviour towards which it will naturally gravitate (Systems Innovation 2020c).

It is critical to appreciate that there does not always have to be an attractor. Whether there is or not is an empirical question⁵.

Here are two versions of the same question.

- What outcome will the programme have?
- What attractor space describes the programme's outcome?

The 'attractor version' leads to inquiry that does not fall naturally out of the 'outcome version'.

- Conceptualizing outcome as a value within an attractor leads to curiosity about the range of values the outcome can take (boundaries of the attractor) and what effort is needed to effect a change

⁵ Understanding whether systems converge to an equilibrium or diverge in unpredictable ways is a major theme in the field of complexity. Two dynamics drive unpredictability. The first is sensitive dependence. The second is that a large number of low-probability events may cause a major disruption in a system. For a dramatic example, see Rumsfeld (2001). For an analysis of the high probability of at least one member of a set of low-probability events occurring, see Taleb (2010). This kind of behaviour is one of the main reasons that evaluators must pay attention to unintended consequences (Morell 2010).

from one set of outcome values to another (topography of the attractor, also known as sustainability, also known as resistance to change).

- Attractors provide a way to kick understanding of programme outcome up a level of abstraction and thus provide insightful comparison between seemingly dissimilar programmes. This is because similar outcome attractor spaces for seemingly dissimilar programmes raises suspicion that maybe those programmes are not so different after all.

As a simple illustration, imagine evaluating a programme designed to increase cooperation between a regulatory agency and industry as a means of improving safety. We know that neither enforcement nor cooperation alone is sufficient to ensure safety (Sparrow 2000). We also know that high-profile accidents push agencies to become more punitive. As a result, the behaviour of regulatory agencies can be visualized as a pendulum that swings over time between excessive cooperation and excessive enforcement. What does this dynamic mean for understanding sustainability in a programme that has successfully improved safety by increasing cooperation? It means that the more successful the programme, the more likely the agency is to reach the 'swing back' point. Note that this scenario has said nothing about which regulatory agency is involved or about any of the details of the safety programme. Rather, it describes the attractor shape for many different organizations and programmes. It allows us to consider similarities among many settings that exhibit that attractor.

Emergence

The whole is different than the sum of its parts. This truism has special meaning in complexity.

Imagine a cylinder in the internal combustion engine of an automobile. I can explain what a cylinder is, how it is constructed, how it fits into an internal combustion engine and so on. Yes, the automobile is different from the sum of its parts, but the uniqueness of the cylinder in the system called an 'automobile' remains. The same holds for organs in a human body or a graphics card in a computer.

Now think of a beehive or a traffic jam or an economy or the vitality of living in a dense urban area. It is *impossible* to explain a beehive in terms of the behaviour of each bee. It is *impossible* to understand a traffic jam in terms of the velocity of each car. It is *impossible* to understand an economy

by breaking it down into the actions of each person and firm that makes up the economy. It is *impossible* to explain urban vitality by analysing the behaviour of each person living in a city. In all these examples, the whole is different from the sum of its parts in the sense that the parts lose their unique identity. When you see that, you see emergent behaviour.

Emergence touches on the question of what should be measured. It is natural to think of the consequences of interventions as being made up of constituent parts, each of which should be measured, but if what matters is the emergent property of many interacting parts, it may be difficult, or even impossible, to conceptualize an outcome in terms of the aggregate consequences of small achievements.

Phase Transition

Phase transitions are about qualitative change that results from small quantitative change.

A phase transition may be defined as some smooth, small change in a quantitative input variable that results in an abrupt qualitative change in the system's overall state. The transition of ice to steam is one example of a phase transition (Systems Innovation 2020d).

Although the term 'phase transition' has its roots in the chemical and physical properties of matter, it can also be applied to human-centric contexts, as for instance the brief time it took the Republican Party in the United States to transform itself from a long history of pro-free trade, pro-immigration, internationalist inclinations to a U.S.-centric political philosophy as Donald Trump rose to prominence and position. Imagine the methodological and analytical differences in evaluating two different models – one that hypothesized quantitative change in the magnitude of an outcome and one that hypothesized a qualitative change that resulted from a small change in an outcome's magnitude.

Self-Organization

Self-organization is a process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system's components are executed using local information, without reference to the global pattern (Santa Fe Institute 2020b).

The key insight in this definition is that a system can form a pattern without 'instruction' from the outside world. This does not mean that external events cannot perturb the system. It does mean that outside events do not control the system.

The possibility that self-organization is present has implications for sustainability and for its inverse – resistance to change. Meaningful evaluation questions include: Is the potential for self-organization present? Is self-organization operating? If a system is perturbed, how long does it take to evolve back into equilibrium? Is self-organization desirable?

Sensitive Dependence on Initial Conditions

Most of us have been schooled to worship at the altar of the general linear model. We have been taught to think in terms of groups – their means, variances and distribution shapes. Everything we do is focused on eliminating the influence of individual data points. We scan for outliers. We make sure our samples are representative of carefully defined sets. We endeavour to keep our variances tight. We base inference on the belief that error across observations will sum to zero. Sensitive dependence, which is a critical construct in complexity, offers a complementary analytical lens – one in which local variation can affect the long-term evolutionary direction of the whole system.

A system's sensitivity to initial conditions refers to the role that the starting configuration of that system plays in determining the subsequent states of that system. When this sensitivity is high, slight changes to starting conditions will lead to significantly different conditions in the future (Santa Fe Institute 2020c).

[Sensitive dependence] refers to the idea that current and future states, actions, or decisions depend on the sequence of states, actions, or decisions that preceded them – namely their (typically temporal) path. For example, the very first fold of a piece of origami paper will determine which final shapes are possible; origami is therefore a path dependent art (Santa Fe Institute 2020d).

Because of 'sensitive dependence', a system's overall behaviour can be understood in terms of how small changes within the system influence long-term trajectories as systems evolve over time. Because of sensitive dependence, it cannot be assumed that a sequence of relationships that exist at one point in time will repeat. Thus, although a causal path can be traced in retrospect, knowing that says little about where the path will lead next.

Combining Complexity Constructs to Explain Outcomes

The previous section addressed individual complex behaviours. Here I will illustrate how these behaviours can cluster to produce an intellectual orientation to pattern, predictability and how change happens.

Stigmergy and self-organization convey a sense that elaborate, seemingly deliberately planned, goal-oriented behaviour need not have central direction. One implication is that programme theories based on deliberate planning may be incorrect portrayals of how coordination takes place. A second implication is that, because theory guides methodology, evaluation will not provide data on the coordination process at play.

Phase transitions and emergence convey a sense that qualitative change can take place in constructs that have quantitative identities. The notion of phase transitions implies that conditions can remain little changed over an extended period and then change suddenly to qualitatively different states, and that of emergence implies that parts of a system lose their identity. Before emergence, it makes sense to observe and measure constituent parts. After emergence, the identities of those parts lose their meaning.

Complexity-Based Explanation

Much of the discussion so far has inclined heavily in the direction of instrumental action. If I know that emergence is happening, I should measure at the aggregate level. If I can identify an attractor, I should use the knowledge to assess resistance to change. And so on. There is a 'predictive' sensibility: 'If I implement this programme, what will happen?'

Evaluation is steeped in this predictive mindset. After all, the whole field is based on the belief that social science can give planners guidance⁶. Our work is technological, not scientific (Morell 1979). 'The aim of technology is to be effective rather than true, and this makes it very different from science' (Jarvie 1983). But what happens when the technology fails, when the predictive ability of evaluation fails to provide guidance to decision makers? Then the need arises to delve into explanation, to understand the science of why events occur (Feibleman 1983). When that need arises, complexity provides a productive framework.

⁶ Donald Campbell's (1991) classic piece on this topic is always worth reading.

I am not arguing that all evaluation of transformation should be based on complexity. I am only arguing that 'all models are wrong, but some models are useful' (Box 1979) and that complexity-based models are useful when transformation is being evaluated. What would a theory of transformation to green energy tell us if we invoked a complexity framework?

What Is Transformation?

An intuitive understanding of transformation is as a transition to a new normal, a default set of conditions that shape how we live⁷. Here are some examples.

- Wood to coal
- Animal to steam power
- Mercantilism to capitalism
- Horse to horseless carriage
- Long-distance fast communication, starting with the telegraph
- Mechanized transportation, starting with railroads and steamships
- The nation-state as a unit of relationships among geopolitical entities
- Income taxes as a legitimate way (at least in the United States) for a government to raise revenue

There were times before these new normals, for example, when people thought that a national surplus defined a nation's wealth, when it was inconceivable that a human could move 50 miles per hour and when information took weeks to move over long distances. What changed?

Let us take the example of the transition from wood to coal in England between the 17th and 19th centuries (Allen 2013; Rhodes 2018). (Yes, it did take a long time.) What needed to be present to effect this change? Steam power was available to drive engines to keep mines dry. Heating demand due to urban density denuded local forests. Patent law and the ratio of

⁷ 'Default' is the operative word that makes 'transformation' different from 'sustainability'. One can think of this in terms of system maintenance. Does energy have to be put into the system to maintain it, or are equilibrium and self-organizational dynamics at play? There is an extensive literature on evaluating sustainability (Julnes 2019). To understand transformation, the concept of default conditions, and the reasons they may or may not arise, needs more attention than it gets.

labour to capital made invention appealing. The building boom in London was conducive to developing new chimney designs. And much else besides.

It is important to identify each of these factors, to assess their behaviour and to determine their interactions, but another useful perspective is to view these changes as a transition from one attractor regime to another, from an equilibrium condition that favoured wood to an equilibrium condition that favoured coal. Within each attractor, the self-organizing capacity of activity within the attractor would counteract any force that perturbed the attractor. That is a perspective that leads to speculation about the shape and depth of the attractor and raises questions that would not arise with traditional evaluation reasoning.

Why would it lead to different strategies? Because it would affect our theories of change. A complexity argument would claim that, within the attractor, it *may* be possible to identify all the relevant components but that it is *impossible* to understand the attractor in terms of relationships between each of those components. Why? Because the equilibrium condition that defines the attractor is an emergent phenomenon. It may be possible to know what the parts are, but it is not possible to identify the specific role of each part. Moreover, if we believe in sensitive dependence, we believe that, each time the attractor is perturbed, the self-organization dynamic might be different. All we can say is that the attractor is deep enough relative to self-organization capacity that, when the attractor is perturbed, it returns to its equilibrium condition.

Defining the Outcome: What Is Green Energy Transformation?

The complexity view tells us that, if there is a transition to green energy, many different factors must come together, but that we do not know (and probably cannot know) what they all are and that, whatever they are, they can come about in different combinations. How to evaluate a scenario like this?

A good place to begin is by defining the desired outcome by making as informed and data-based a guess as possible to answer the question: How much use of green technology is needed to make it the default choice for the foreseeable future? Here is an example of what might work as a definition of the transformation: 'We know that transformation has happened if, in geopolitical boundary X, approximately 80 per cent of energy use comes from green sources and has remained at approximately that level

for five years'. I like this form because it includes different dimensions of whatever attractor space constitutes a green energy new normal, as shown in table 16.2.

Table 16.2 Elements of the Green Energy Attractor

Element	Description
Geopolitical boundaries	The existence of geopolitical boundaries implies a reasonably large geographic area. It is a proxy for availability of equipment, businesses and expertise to install and maintain systems; cost; political consensus; the reach of regulation and peer pressure.
Level of energy use	'Eighty per cent' is a level of energy use that would be truly different from the old way of doing things. It could only come about from a profound change in energy sources and their supporting constituents.
Time	'Five years' acknowledges that an indicator of profound change requires assurance that the construct it reflects will endure over time. It indicates that the attractor is stable.
Imprecision	The definition acknowledges that there is a range that defines the boundaries of the attractor.
Measurement	Everything in the definition can be measured.

What if the definition turned out to be wrong? That would be OK. It would mean that evaluation revealed a problem with stakeholders' programme theory and, it is hoped, guidance for correcting the theory.

Complexity-Based Models

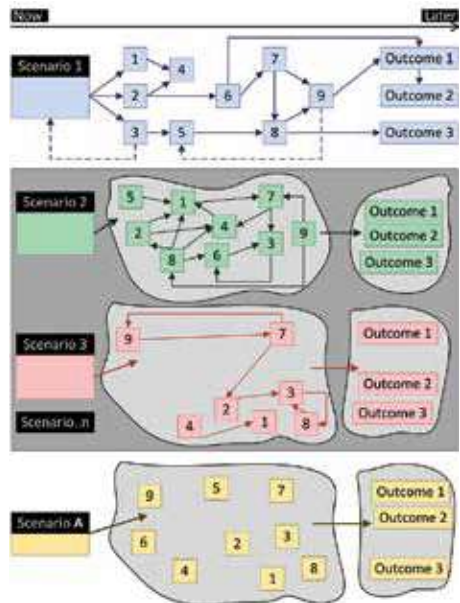
Empirical inquiry requires an exercise in wilful ignorance (Weisberg 2014) because there will always be relationships we care about that are enmeshed in a multitude of relationships that complicate and obscure what we want to know. Therefore, any research enterprise, evaluation included, requires a model, a simplified version of reality that specifies the relationships we care about. Any model we use will be wrong, but some will be useful (Box 1979). Complexity-based models are wrong but useful in ways that our traditional models are not.

Comparing Complexity-Based and Non-Complexity Based Models

Compare the scenarios in figure 16.1.

- *Scenario 1.* Scenario 1 assumes that we know enough to specify all outcomes and relationships among their antecedents.
- *Scenarios 2, 3...n.* These scenarios (grey field in figure 16.1) depict different possible complex relationships between the programme and its outcomes. What is the message in these scenarios? (1) There is a connection between programme and outcome. (2) There is an exceedingly large number of paths that can elicit the desired outcomes. (3) Because of sensitive dependence, we cannot predict the precise path between programme action and desired outcomes. (4) Not all the known relevant factors must be equally important during each pass through the system. (5) No single intermediate factor leads directly to any of the desired outcomes. Rather, outcome stems from the emergent effect of all the networked intermediate elements. So, when designing an evaluation, which configuration should we pick? None. Why? Because whichever we choose, that configuration may be different in the future.

Figure 16.1 Traditional and Complex Models for the Same Programme



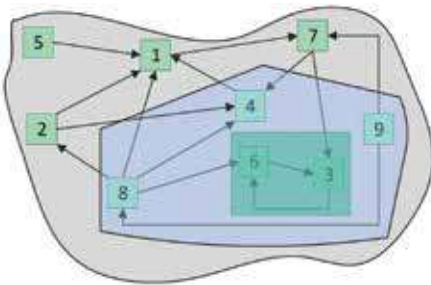
- *Scenario A.* Scenario A has a very simple logic: Do a lot here, and something will happen there. In a world driven by complex behaviours, this logic makes sense. I do not mean to imply that we can pick the internal elements of the model at random. It is important to specify categories that need to be included (e.g. economic conditions, technological capabilities, regulatory structures, culture). After all, we have domain knowledge based on experience, research and theory.

Possibilities for Prediction

I overstated the case and left the impression that highly specified models cannot be predictive. What is true is that the broader the scope of a model, the greater the likelihood that complex behaviours will replace the role of specific relationships. Put differently, a model might be everywhere correct locally but incorrect globally.

Figure 16.2 illustrates this point. It is drawn from scenario 2 in figure 16.1. The simplest region (green rectangle) contains two elements connected with a single feedback loop and two direct connections with the outside. The next larger region (blue) contains five elements. It also contains nested feedback loops and three direct connections with the outside. Finally, there is the entirety of scenario 2.

Figure 16.2 Global/Local Model Correctness



I do not know how to quantify degrees of complexity, but it seems reasonable to subject the green region to a traditional evaluation. I am not sure I would do it for the blue region, but I could be convinced. I know I would not accept that tactic for the entire model.

It is also important to keep in mind that the argument above is about predicting a causal path rather than tracing a causal path that has already occurred. Nor is it about identifying all the elements in the model. When sensitive dependence is operating, what cannot be predicted for any given path through the model is which elements will be active and how they will relate to each other. Once the model runs its course, all can be identified.

A Meta-Theory of Transformation to Green Energy

Existing theories of transformation clearly engage complexity. Some engage complexity implicitly. Complex behaviour is contained within the model, but there is no explicit mention of complexity. Other theories explicitly draw on complexity. I will give an example of each and then make the case that theories of transformation should draw from complexity science in a systematic fashion.

Theories of Transformation That Do Not Explicitly Refer to the Field of Complexity

Reed and Jordan (2007) developed a systems theory for the U.S. Department of Energy's Energy Efficiency and Renewable Energy (EERE) programme. They confronted a classic complex system problem. EERE has the long-term goal of engendering a regime of efficient renewable energy. In doing so, it runs many discrete programmes that emanate from many different cubbyholes within the Department of Energy, all of which have different short- and intermediate-term goals and separate theories. With respect to the long-term goals, the theories employ the well-known logic: implement programme → accomplish short-term goals → magic happens → achieve long-term goals.

Reed and Jordan's proposal was that all the diverse programmes conduct evaluation based on Rogers' (2003) theory of innovation. That theory's constituent parts are applicable to a wide range of settings, making evaluation findings comparable across diverse contexts. Individual programmes may still need their own unique objectives, but by invoking Rogers, the diverse programmes can also share goals. Because of this commonality, the strengths and weaknesses of separate programmes can be compared. Reed and Jordan do not discuss their efforts in terms of complexity, but it is clear that stigmergy, emergence and sensitive dependence can provide complexity-inspired explanations about pattern, predictability and how change happens.

Stigmergy

Recall that stigmergy is a process in which a plan is embedded in the history of activity that independent actors encountered (Theraulaz and Bonabeau 1999). Now consider EERE's dilemma. Their various programmes have a common long-term goal and different short-term goals and are embedded in a bureaucracy that makes tight coordination difficult and counterproductive (Morell 2018).

EERE can use knowledge of common goals to make organization-wide decisions, but something else is also going on. Reed and Jordan have devised a mechanism that changes the information environment such that each programme can make independent self-interested decisions informed by what its surrounding programmes have, and have not, done. This is stigmergic change.

Implement programme → accomplish short term goals → magic happens → achieve long-term goals. Emergence and sensitive dependence explain the magic. Consider the EERE scenario in light of figure 16.1.

Many different programmes inhabit the same ecosystem. At any single decision point, an individual programme may make a decision that links to other programmes. But what linkages? And to which programmes? Those decisions will be based on judgments made at unique points in time based on demands of the moment. At other times, or under different perceived conditions, decisions will lead to a different set of linkages. Where does this leave evaluation of EERE?

- Because of previous research, we know what elements must be included. Evaluation can determine which ones have.
- Because of the common Rogers-based goals, some form of coordination might take place as each programme makes its own decisions. Evaluation can tell us the whethers, whats, whys and hows of that coordination.
- Complexity-informed programme theory tells us that:
 - Because of sensitive dependence, the chain of coordination relationships cannot be specified in advance or relied upon to endure over repeated planning cycles.
 - Success may be a function of the amount of coordination but not of what specific coordination took place.
 - Because network linkages are involved, success may be an emergent function of the linkages; that is, the overall effect cannot be explained in terms of the unique identity of each of its constituent parts.

Complexity explains the 'magic'. It is not magic at all. It just seems like magic because complex behaviour may not conform to our common sense.

Theories of Transformation That Explicitly Refer to the Field of Complexity

Considerable effort is being made to draw on complexity when developing theories of transformation. What we need is to enrich and systematize this line of thinking. Three examples illustrate how current thinking about theories of transformation have drawn on complexity.

Example 1. Zazueta (2017) has proposed a theory of change that draws heavily on networking (figure 16.3). He identifies adaptive learning, feedback

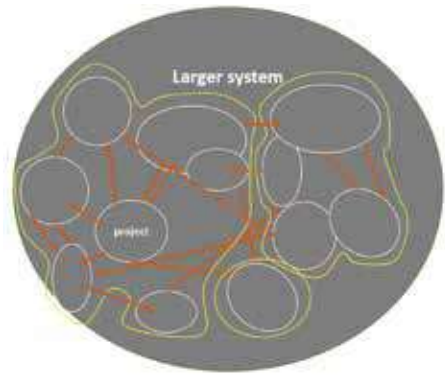
and emergence as behaviours of networks. He also specifies that ‘agents’ are operating and notes the importance of domains and scales of space and time. The graphic implies that there are nodes and edges, but precisely what they are and how they are connected is left undefined.

Example 2. Figure 16.4 illustrates the theory of transformational change that the SDG Transformation Forum (2020) proposed. It relies on feedback loops and networks but acknowledges that specific elements of success are unknown, hence the unlabelled network nodes and the question marks that, presumably, are there to indicate uncertainty about network edges.

Example 3. Figure 16.5 is an adaptation of a model that Fisher and Roehrer (2020) developed to understand progress towards transformation. Individual elements (incremental inputs on the left side of figure) undergo a network development process that transforms them into transformational elements on the right (e.g. projects and portfolios).

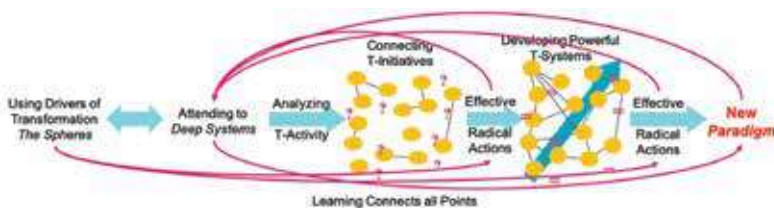
All three examples specifically identify network behaviour as crucial to transformation. All three acknowledge two domains of uncertainty – the specific identity of nodes (relevant variables) and the causal relationships among these nodes.

Figure 16.3 Theory for Transformational Development

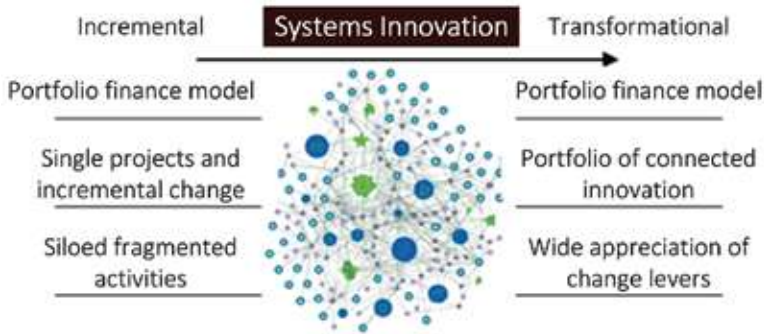


Source: Zazueta (2017).

Figure 16.4 Theory of Transformational Change



Source: 2020 SDG Transformations Forum.

Figure 16.5 Model of Progress Towards Transformation

Source: Adapted from Fisher and Roehrer (2020).

Extending the Application of Complexity in Devising and Using Theories of Transformation

Table 16.3 identifies the complex behaviours contained in these theories.

The complex behaviours implicit in existing theories of transformation should be made explicit and considered in a deliberate manner. A meta-theory of transformation is useful for facilitating such deliberate consideration.

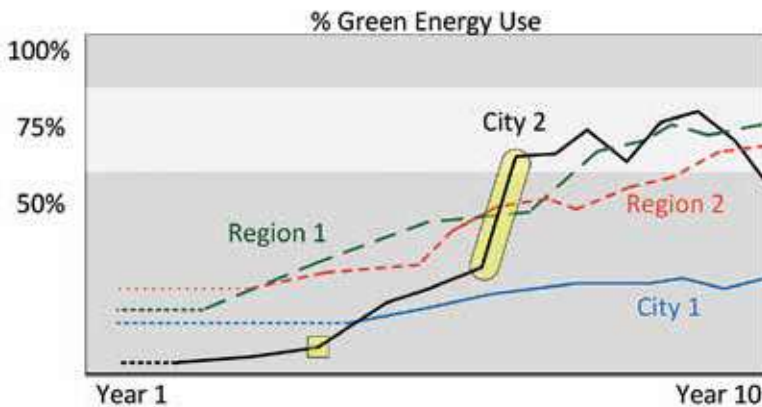
To produce a theory of transformation, it is necessary to begin by defining the outcome, in this case the criteria in table 16.2: geographical boundaries, level of green energy use, geographical spread, range not point estimates and making sure that it is all measurable.

By defining outcomes this way, it will be possible to produce data as depicted (in the entirely fictional scenario) shown in figure 16.6. In the figure, colours represent geographical entities, dashed lines represent regions, solid lines represent cities and straight dotted lines show the time in each location before which any change might be expected. How might a complexity perspective interpret this data?

- For there to be a 'new normal', geographical spread matters because geography is a proxy for availability of equipment, businesses and expertise to install and maintain systems; cost; political consensus; the reach of regulation and peer pressure. Figure 16.6 shows which locations changed and when the changes took place.

Table 16.3 Complex Behaviours Implicit in Existing Theories of Transformation

Complex behaviour	Manifestation in theories of transformation
Emergence	Emergence can account for reasons to avoid explaining transformation in terms of linear combinations of discrete variables.
Phase transition	Phase transitions are common as edges grow in a network.
Attractor behaviour	Attractors allow for the fact that, despite the uncertainties of sensitive dependence, there are circumstances under which, if enough activities are done well, specific outcomes can be expected.
Sensitive dependence	Sensitive dependence implies that, even when a causal chain can be determined in retrospect, that same causal chain may not operate in the future.
Stigmergic and self-organizing phenomena	Stigmergic and self-organizing phenomena may drive activity in the direction of organized change even absent tight central coordination.

Figure 16.6 Illustration of Evaluation Data, Promotion of Green Energy Technology

If the data were paired with a map, evaluators would have a solid appreciation of how infrastructure support evolves.

- Complexity posits that, even if change is defined as space within an attractor, there is still the question of the topography of the attractor – how well can self-organizing forces 'hold' values within the attractor? Eyeballing the data suggests that the attractor seems stable for larger geographical areas (regions) even if it may not be stable for smaller areas (cities).
- The definition of success stated a range for percentage of green energy. Three of the four regions made it into that range but only into the bottom of the range, and one of those almost fell out. Perhaps the natural range for green energy use under the interventions implemented and in the environment in which they were implemented is lower than what was expected? This may suggest a change in programme theory or an adjustment in our understanding of realistic outcomes.
- To say that approximately 80 per cent green energy use is a new normal is to say that it is qualitatively different from lower percentages. This may be the case because all of the factors that affect energy use come together in a networked fashion to yield an emergent condition in which component parts lose their identity. Is our hypothesis correct that emergence takes place at approximately 80 per cent green energy use?
- Complexity tells us that phase shift behaviour is possible. It does not tell us that there *must* be such change or that the new normal cannot happen incrementally. See the yellow oval for city 2. It seems as if a phase shift may have taken place. Incremental change seems to be the case in the other scenarios.

What complex behaviours would have to be built into the evaluation to allow us to interpret the data in complex terms? The answer is summarized in table 16.4.

In addition to the implications of the specific complex behaviours described above, a complexity perspective constitutes a style of reasoning. Table 16.5 gives some examples. All of these examples speak to the themes in complexity that constitute the columns in table 16.1 – what pattern we can expect, what we can and cannot predict, how change happens.

Finally, drawing on complexity can help when efforts at transformation fail because the process of transformation is a complex system, and

Table 16.4 Complexity as It Applies to Theories of Transformation

Emergence	<ul style="list-style-type: none"> ▪ Does the model identify what the emergent outcome is? ▪ Does the methodology consider the emergent behaviour as its own variable? ▪ Does the theory of change recognize the importance of individual elements without assuming that the consequences of those elements can be 'added up'? ▪ Does the model reflect when an emergent change will appear or for how long into the future the change will persist?
Phase transition	<ul style="list-style-type: none"> ▪ Does the theory postulate a non-linear change in which little happens for an extended period of time? ▪ Do the theory and methodology (not to mention stakeholder expectations) acknowledge that the concept of 'intermediate stages of transformation' may not have much meaning? ▪ Does the model recognize timing, that is, does it identify a window within when the change can be expected?
Attractor	<ul style="list-style-type: none"> ▪ Does the theory acknowledge that transformation may be defined as an attractor that can be explained as a condition in which self-organization resists changes to the status quo? ▪ Has any thought been given to how deep that attractor is, that is, how resistant the transformation state is to outside shocks? ▪ Does the methodology consider the stability of the attractor? Put in other terms, if the model predicts the appearance of an outcome attractor, for how long will that prediction remain accurate?
Sensitive dependence	<ul style="list-style-type: none"> ▪ Does the theory specify relationships between discrete elements, or does it recognize the possibility of sensitive dependence, a condition in which multiple unpredictable chains of causation may lead to the same result? ▪ How does the evaluation engage this possibility in terms of metrics that specify what needs to be measured and a methodology that provides the logic of data interpretation?
Stigmergy	<ul style="list-style-type: none"> ▪ Does the theory explicitly consider coordination among the actors involved in transformation activities? ▪ If so, does the theory consider the possibility of stigmergic processes in which independent choices are influenced to work towards a specific goal?

Table 16.5 Examples of Reasoning that Derives from Combining Complex Behaviours

Complex behaviour	Implication for understanding change
Emergence and phase transition	Combine to convey a sense that smooth incremental change is not typical behaviour
Sensitive dependence and attractors	Combine to convey a sense that clearly specifiable patterns should not be expected
Stigmergy, attractors and sensitive dependence	Combine to convey a sense that, even without high levels of process control, certain outcomes can be expected

therefore, the science of complexity is needed to explain success and failure when the theory of action (the technology) of effecting change fails.

Is There a Recipe for Applying Complexity to Evaluation?

No. There is no recipe. What I can offer is a set of questions to ask, put in a sensible order.

- *What are the characteristics of the desired state?* These need to be defined in terms of multiple measurable elements and levels of imprecision.
- *Is the desired state a 'new normal'?* Will the desired condition be the default, or will it need energy to sustain it?
- *Is the desired state qualitatively different or just more (or less) of what went before?* Is it more like an economy or a traffic jam or an overall health measure consisting of different kinds of health improvements?
- *What does the outcome chain look like?* Begin with a traditional deterministic model. Then ask: Is this the only path through the model's parts that will lead to the desired state? Are there other elements that *might* be operating even if I don't see exactly how they fit? Is it likely that elements I cannot foresee might become relevant? Might small local changes affect the entire path through the model?
- *How do the coordination mechanisms work?* Question whether direction is imposed or emerges from independent action.

Acknowledgements

Many people provided insightful critique on this paper. Thanks go to Marie H el ene Adrien, James Altschuld, Rob D. van den Berg, Neil Bird, Cristina Magro, Apollo Nkwake and Beverly Parsons.

References

- Allen, Robert C. 2013. Energy transitions in history: The shift to coal. In: *Energy transitions in history: Global cases of continuity and change*, (ed.) Richard W. Unger, 11–15. Munich: Rachel Carson Center for Environment and Society.
- Bamberger, Michael, Leonardus Vaessen and Estelle Raimondo (eds.). 2015. *Dealing with complexity in development evaluation: A practical approach*. Thousand Oaks CA: Sage.
- Boehnert, Joanna. 2020. The visual representation of complexity. <https://tinyurl.com/48u5r8e3>
- Box, G.E. P. 1979. Robustness in the strategy of scientific model building. In: *Robustness in statistics*, (eds.) R. L. Launer and G. N. Wilkinson, 201–236. Cambridge, MA: Academic Press.
- Campbell, Donald T. 1991. Methods for the experimenting society. *Evaluation Practice* 12 (3): 223–260.
- Castellani, Brian. 2014. Brian Castellani on the complexity sciences. In: *Theory, Culture, & Society* <https://tinyurl.com/yatnw78>
- Castellani, Brian. 2009. Map of complexity sciences. In: *Art & Science Factory*. https://www.art-sciencefactory.com/complexity-map_feb09.html
- Feibleman, J. K. 1983. Pure science, applied science and technology: An attempt at definitions. In: *Philosophy and technology*, (eds.) C. Mitcham and R. Mackey, 33–41. New York: Free Press.
- Fisher, Susannah, and Christine Roehrer. 2020. Understanding progress towards transformation, in time. EIT Climate-KIC. Amsterdam.
- Jarvie, I. C. 1983. Technology and the structure of knowledge. In: *Philosophy and technology*, (eds.) C. Mitcham and R. Mackey, 54–61. New York: Free Press.
- Julnes, George (ed). 2019. *Evaluating sustainability: Evaluative support for managing processes in the public interest*. New York: Jossey-Bass.
- Marion, Russ. 1999. *The edge of organization: Chaos and complexity theories of formal social systems*. Thousand Oaks, CA: Sage.
- Morell, Jonathan. Systematic iteration between model and methodology: A proposed approach to evaluating unintended consequences. *Evaluation and Program Planning* 68: 243–252.
- Morell, Jonathan. 1979. *Program evaluation as social research*. New York: Pergamon Press.

- Morell, Jonathan. 2010. *Evaluation in the face of uncertainty: Anticipating surprise and responding to the inevitable*. New York: Guilford.
- Morell, Jonathan. 2017. Complex behavior in program design and evaluation: Toward furthering a dialogue between program designers and evaluators. Presented at the annual meeting of the Associazione Italiana di Valutazione. Padova, Italy.
- Morell, Jonathan. 2019. Ignoring complexity can make sense – Part 3 of a 10-part series on how complexity can produce better insight on what programs do, and why. <https://tinyurl.com/5dw26kru>
- New England Complex Systems Institute. 2020. New England Complex Systems Institute portal. <https://necsi.edu/>
- Niiniluoto, Ilkka. 2019. Scientific progress. In: *The Stanford encyclopedia of philosophy*, (ed.) Edward N. Zalta. Palo Alto, CA: Metaphysics Research Lab, Stanford University.
- Parunak, H. Van Dyke. 2006. A survey of environments and mechanisms for human-human stigmergy. In: *Proceedings of E4mas 2005*, (eds.) D. Weyns, F. Michel and H.V.D. Parunak, 163–186. New York: Springer.
- Phelan, Steven E. 2001. What is complexity science, really? *Emergence* 3 (1): 120–136.
- Reed, John H., and Gretchen Jordan. 2007. Using systems theory and logic models to define integrated outcomes and performance measures in multi-program settings. *Research Evaluation* 16 (3): 169–181.
- Rhodes, Richard. 2018. *Energy: A human history*. New York: Simon and Schuster,.
- Rogers, Everett. 2003. *Diffusion of innovations*, 5th ed. New York: Free Press.
- Rumsfeld, Donald. 2001. *Predicting the future*. Washington, DC: Department of Defense.
- Santa Fe Institute. 2020a. Complexity explorer. <https://www.complexityexplorer.org/home>
- Santa Fe Institute. 2020b. Self-organization. <https://tinyurl.com/4nwzd9u5>
- Santa Fe Institute. 2020c. Sensitive dependence on initial conditions. <https://tinyurl.com/2bdvp82w>
- Santa Fe Institute. 2020d. Path dependence. <https://tinyurl.com/r52r4hfs>
- SDG Transformations Forum. 2020. Our theory of transformational change. <https://transformationsforum.net/totc/>
- Shmueli, Galit. 2011. To explain or to predict? *Statistical Science* 25 (3): 289–310.
- Sparrow, Malcolm K. 2000. *The regulatory craft: Controlling risks, solving problems, and managing compliance*. Washington, DC: Brookings Institution.
- Systems Innovation. 2020a. Systems innovation portal. <https://systemsinnovation.io/>
- Systems Innovation. 2020b. Systems innovation glossary – attractor. <https://systemsinnovation.io/glossary/attractor/>
- Systems Innovation. 2020c. Social attractors & chaos theory. <https://www.youtube.com/watch?v=C9Xl6BNu-kM>
- Systems Innovation. 2020d. Phase transition definition. <https://tinyurl.com/j5fc7m6w>

- Taleb, Nassim Nicholas. 2010. *The black swan*, 2nd ed. The impact of the highly improbable. New York: Random House.
- Theraulaz, Guy, and Eric Bonabeau. 1999. A brief history of stigmergy. *Artificial Life* 5: 97–116.
- Tyrrel, Lavinia. 2019. Theory of change and theory of action: What's the difference and why does it matter? <https://tinyurl.com/bejpn33f>
- Weisberg, Herbert I. 2014. *Willful ignorance: The mismeasure of uncertainty*. New York: Wiley.
- Zazueta, Aaron. 2017. *Principles for the development of integrated transformational projects in climate change and chemicals & waste*. Washington, DC: Scientific and Technical Advisory Panel of the Global Environment Facility.



CHAPTER 17

Of Portals and Paradigms: Evaluation, Systems Thinking and the Pandemic

CRISTINA MAGRO AND ROB D. VAN DEN BERG

Abstract. This chapter proposes an exercise with systems thinking, taking the COVID-19 pandemic as a platform for learning, to illustrate the kind of reasoning, language and narrative that will help evaluators focus on key questions and approaches that are adequate. With this, the authors hope to help strengthen and spread the paradigm of systems thinking in evaluation. The authors argue that all social, economic, environmental, cultural and cognitive contexts are here to support evaluators dealing with systems thinking. After a presentation of systems thinking phenomena relevant for the exercise, the chapter takes readers on a journey through a broad, interrelated view of the experience with the pandemic and presents quick takeaways and consequences for evaluators and evaluation. Throughout the journey, the habits of a systems thinker are followed to gain insights and a natural flow of reasoning in systems terms.

Historically, pandemics have forced humans to break with the past and imagine their world anew. This one is no different. It is a portal, a gateway between one world and the next.

We can choose to walk through it, dragging the carcasses of our prejudice and hatred, our avarice, our data banks and dead ideas, our dead rivers and smoky skies behind us. Or we can walk through lightly, with little luggage, ready to imagine another world. And ready to fight for it. –Arundhati Roy (2020)

A Journey with Systems

In *Systems Evaluations for Transformational Change: Challenges and Opportunities*, 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 (Magro and Van den Berg 2019, 131).

Less than one year after this publication, evaluators encountered an unexpected opportunity to experiment with systems thinking, to realize how tangible it can be and to grant unprecedented relevance to the above claims – following Arundhati Roy – to break with the past and imagine a world anew, crossing a gateway between one world and the next.

Working with paradigm changes in science, anthropologist and sociologist of science Bruno Latour (1987; 1988; Latour and Woolgar 1986) emphasizes the value of observation and experience, history and context in the development, acceptance and consolidation of ideas. He focuses on the paramount role of context and emphasizes the non-linearity of scientific development. In one of his books, *The Pasteurization of France*, Latour (1988) argues that the success of Louis Pasteur's work – related to the spread of microorganisms and epidemics such as cholera – needs to be understood within the actual historical convergence of contexts. In Pasteur's time, these included the public hygiene movement, the medical profession and colonial interests. Without the conjunction of these elements, he argues, his scientific work would not have fully developed and, even if published, would

probably have remained unnoticed. The analysis he proposes is meaningful here, as we consider that the current context is extensively favourable for the evaluator's effective engagement in systems thinking and that the pandemic creates space for reflection about and experimentation with the way of perceiving things, the reasoning, language and narrative involved in systems thinking.

Systems thinking has steadily grown in evaluation over the last two decades, especially related to the Sustainable Development Goals (SDGs), climate and the environment, agriculture, gender and social systems. The SDGs present a shared vision of aspirations related to peace, people, planet, prosperity and partnership. Agenda 2030, in which the SDGs are embedded, calls for a transformation of our world. An immediate consequence of addressing such an ambitious, holistic and aspirational agenda, focusing on transforming social, economic and environmental systems, is the need to provide a way of reasoning that can treat it appropriately.

Blue Marble Evaluation: Premises and Principles (Patton 2020a) strengthens, in the compelling call it conveys, that being fluent in systems thinking is critical to addressing the SDGs and Blue Marble issues and to transforming evaluation itself. Patton joins a broad community alerting the world that the massive threats currently being experienced can no longer be neglected, nor can we postpone the quest for novel ways of thinking to approach them. In his blog *Evaluation Implications of the Coronavirus Global Health Pandemic Emergency*, Patton (2020b) recommends:

7. Engage in systems thinking. If you have been putting off bringing systems thinking to your evaluations, *now is the time*. If you've already been bringing systems thinking to your work, *now is the time* to go deeper and demonstrate to those you work with the relevance and importance of thinking systemically about what is happening. Public health, community health, national health, global health, your family's health, and your personal health are all connected. This is micro to macro, and macro to micro, systems thinking. The state of public health is connected to the economy, the financial system, politics at every level, social well-being, cultural perspectives, educational inequities, social and economic disparities, public policy decisions, and evaluation. *Practice seeing the interconnections and their implications for your work, your evaluations, and your life.* Celebrate the initiatives of young people worldwide to build a more sustainable and equitable future [emphasis added].

We concur with Patton that the time is now to transform the way we think as evaluators, as professionals and as citizens and to dive into what is still often referred to as *the new paradigm*. The context is fully favourable

and, more than that, demands transformational actions. Systems thinking requires sharp, reflexive, critical, dynamic ways of thinking in terms of generative mechanisms that lead us to understand the functioning of things.

This first section – a journey with systems – builds on our previous work (Magro and Van den Berg 2019) exploring conceptual tools to prepare for the journey with systems that comes next. The second, and most extensive, section – a personal journey – explores the Habits of a Systems Thinker¹, taking the pandemic as a platform for learning. On the way, we post ‘road signs’ indicating the habits of a systems thinker that we identify in the narrative to define, enrich and consolidate the reflexive experience proposed. Habits will often be applied simultaneously, but for clarity’s sake, we treat them separately in this chapter. We will not translate this journey into the field of evaluation, because we want to show how systems thinking leads to and requires broader understanding and systemic expression. We want to operate with systems thinking to understand the experience of the pandemic, hoping you will cross this section with the joy of an explorer. Systems thinking offers a palette of visual resources to make explicit the reasoning and the mechanisms and processes involved. In spite of the clear utility of these graphic resources, in this chapter we want to emphasize attitudes and different reasonings that may be fostered to promote the transformation required towards systems thinking and, as a final goal, achievement of the SDGs. In the third and last section – transforming evaluation – we indicate consequences for evaluators and highlight what emerged from our exercise as potentially useful contributions.

We have previously defined systems as

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 [emphasis added] (Magro and Van den Berg 2019, 144).

We claim, in the above quotation, that the functioning of a system is of utmost interest in systems thinking. One of the motivations for systems thinking is precisely to provide a way to deal with dynamic structures and

¹ Waters Center for Systems Thinking (<https://waterscenterst.org/>). Habits 1–14 in this chapter are quotations extracted from the several Waters Center tools updated in 2020. The numbering of the 14 habits is ours and yields internal references only, with no hierarchy implied. For your exercise and delight, you can seek to identify which habits are involved in each situation besides the ones we indicated along the chapter.

historical processes – that is, processes that happen over time. The quotation states that the functioning of a system results in the emergence of phenomena not reducible to the system's components. The system's structure (formed by components and interactions between them) determines its functioning and what emerges from it, which in turn is related to the historical path of interactions with the context, where behaviour is observed. The aphorism 'the structure of a system determines the behaviour of the system' synthesizes this understanding.

Systems function in a context and maintain a permanent flow of interrelations (here also referred to as 'interactions') with it throughout time in such a way that changes in the context can trigger changes in the functioning of the system, and changes in the system can trigger changes in the context. It all depends on the structure of the systems involved, which develop their history – a permanent flow of interactions – in a pairing mode or structural coupling with the context. System and context are always 'adapted' to each other in a complex, dynamic and contingent way. Examining the history of interactions of a system, systems thinkers indicate that a system's behaviour exhibits recurrent patterns over time. The distinction of these two domains of inquiry (system and context) in their permanent flow of interactions is crucial for clear reasoning with a systems perspective. Things become more challenging when the context is formulated in terms of systems, which, in turn, can have their own structures identified. In evaluation, this has been explored in many situations, for example regarding the interaction between social and economic systems and most explicitly between human systems (social, economic) and environmental systems (ecosystems, species systems). At the nexus between environment and development, this has led to the need to recognize environment and development as two evaluands that have different structures, timing, scaling and locations (Rowe 2012; see Uitto 2014 and Uitto, Puri and Van den Berg 2017).

One significant consequence of the way systems function, as above described, is that no system can be informed, instructed or forced, from the outside, to do what it cannot do. To say the least, the outcomes of any intervention vary because of the system's structure (components, their interrelations and patterns over time); because of its structural configuration at the time an intervention is implemented; because of its path of interactions with the context through time and contingent occurrences (e.g. as an outburst of a pandemic); because of the characteristics of the intervention itself and all relevant factors involved, which is then considered part of the context of the system.

Projects, policies and programmes traditionally rely on isolated, immediate, linear and causal relations and on the vision that, if properly designed and implemented, they will make the system move in the planned direction. Taking a systems view, though, evaluators are led to understand that this is not so. The appearance of unpredictable effects can be understood as ordinary signs of how the systems involved cope with the intervention, which may differ from what was planned. Moreover, in view of the demand for transformational changes, changes in one system may trigger changes in interrelated systems, which can account for transformations that are distant in space and time.

The systems thinking here explored is rooted in Humberto Maturana's work on neurophysiology of vision and autopoiesis, later developed as the biology of cognition and language (Maturana and Varela 1973; 1992). This systemic approach entails a complex interactive engagement between what we perceive, what we talk about and what we know, which leads to the collective crafting of a world in all its complexity rather than a breaking up of the world into silos and narrow fields of cognition. *The Habits of a Systems Thinker* (Benson and Marlin 2019), extensively used in this chapter as a transformational approach to learning, problem-solving and understanding the world, also supports this view. 'It's about seeing life in motion, recognizing that the big picture is rarely static, but almost always a web of factors that interact to create patterns and change over time' (Waters Center for Systems Thinking 2021).

Exploring Experience and Building Habits: A Personal Journey

There's nothing more practical than a good theory. –Lewin (1952, 169)

The COVID-19 pandemic has led each of us around the world to experience changes in our lives that have affected the way we interact with others, personally and professionally. Our movements through the air, in water and on the land were affected. Our ordinary routines were turned upside down, and the readiness to engage in everyday activities and pleasures was disturbed. Mundane tasks such as buying food, medication and household supplies;

Habit 1 of a Systems Thinker: Seeks to understand the big picture.

A Systems Thinker focuses on the forest as well as the details of any one tree.

exercising; obtaining medical care and tests; going for a haircut; fixing a leak in the kitchen; having the car washed were categorized into essential and non-essential activities. Only those classified as essential could continue to be performed, under strict rules, redesigned schedules and new hygiene protocols. Constraints of all sorts were experienced to avoid physical proximity with others – no shaking hands; visiting friends and family; going to school, temple, church, restaurants, theatres.

After a while, many no longer complied with the protocols and protested for their rights to come and go, alien to the complexity and unpredictability of the contagion and the disease itself, and the consequences, in the long run, of losing control of looming outbursts. As flexibilization began, protocols were

reviewed according to new observations and experiences. The virus spread rapidly in countries such as Italy, the United States, Brazil and India, especially during the first half of 2020. The second wave emerged in September and October, worse than the first, while, in general, the behaviour of citizens and the government regarding the pandemic remained the same.

At the same time, people acquired new habits and developed new skills, sometimes making enduring transformations in their lives. New talents emerged in the kitchen, in the garden, in households and in families. Some became musicians or painters; others redirected their lives for good. Technology contributed with effective solutions, in a 'tech-celeration' (Standage 2020) that favoured adoption of various technological behaviours.

QR codes appeared in restaurants, bars and cafés in substitution for printed menus and bills, which go from hand to hand and could be a vehicle of contagion. Ticket counters and pass readers activated by hand wave were made available for the same reason. Virtual spaces to keep our minds fresh, to work and to learn increasingly occupied our locked-down lives. Teachers reinvented classrooms and education and, without any previous knowledge, quickly adjusted their practices to virtual spaces. Despite unintended effects that

Habit 2 of a Systems Thinker:

Observes how elements within systems change over time, generating patterns and trends.

A Systems Thinker sees change over time as the dynamics of a system.

Habit 3 of a Systems Thinker: Identifies the circular nature of complex cause and effect relationships.

A Systems Thinker sees the interdependencies in a system and uncovers circular causal connections.

early studies have described, indicating how stressful virtual contacts could be (Sander and Bauman 2020), people progressively diversified their use, including for informal daily situations and to meet with relatives and friends. E-commerce flourished around the world, with more people learning to buy online and more businesses adjusting to the new reality. Telemedicine became common in many parts of the world, modifying the configuration of, for example, psychotherapists' clienteles, potentially allowing for worldwide enrolment. With the appearance of COVID-19, the use of telemedicine and technological novelties and subsequent innovations exploded.

As the West struggled with massive contagion and cruel death statistics, Japan kept the number of COVID-19 cases relatively low through 2020. Explanations of all kinds were offered – including the lifelong habit of the Japanese to protect their mouths and eyes when they are sick, a sense of citizenship and a fear of judgement if they transmitted the virus because they neglected to follow safeguards. In an interview for the journal *Diplomacy*, a Japanese Foreign Policy Forum vehicle, a professor from the Department of Virology of the Tohoku University Graduate School of Medicine, Dr. Oshitani Hitoshi (2020), explained that a systems approach differentiated Japan from Western countries in their disease response. 'Japan's strategy was "to see the forest to understand the whole";' instead of focusing on the trees. The Japanese strategy can be summarized in four points, as follows:

1. *Awareness and early observation, enlarged view, focus on interrelations and on the dynamics of the spread.* Japanese authorities and scientists reacted immediately to the appearance of the first cases and focused on clusters (not on individuals) as potential sources of contamination and outbreaks. China detected the first cases by the end of 2019, and in Japan, 11 individuals with a travel history to China were identified as having the disease from January to early February 2020. At the time, an estimate of several tens up to a hundred infections was made. The wave that emanated from this first cluster, composed of tourists moving around Tokyo, Osaka and Hokkaido sightseeing and congregating with others, was controlled by mid-March. A second wave began in early February through international travellers (often for work or business), with 300 confirmed cases coming from Europe, the United States, Southeast Asia and Egypt and estimates of some 1,000 to 2,000 cases. Because restrictions on circulation were not imposed until

the end of March, infected people moved around the country, resulting in a large outbreak.

Whereas Japan had already had two waves by March, Western countries did not identify a second wave until September to October because they focused on infected individuals, not clusters. Two waves were also observed in Europe but did not lead to focused treatment or specific measures to limit infections.

2. *Focus on the context, patterns, trends of infections and behaviour over time.* Japanese authorities concentrated on identifying and understanding the clusters and their dynamics in the context they moved, tested and monitored the clusters' contexts and tolerated some degree of low transmission rates, allowing them to address the origin of high transmission rates.

Habit 4 of a Systems Thinker: Checks results and changes actions if needed: 'successive approximation'.

A Systems Thinker intentionally gathers information to assess progress before changing actions.

3. *Reliance on scientific understanding of the new virus and the infection processes.* Japan sought to understand the characteristics of this specific virus and its flow of transmission. As long as the authorities could prevent clusters where one infects many, most chains of

Habit 5 of a Systems Thinker: Pays attention to accumulations and their rates of change.

A Systems Thinker clearly identifies elements of the system that accumulate and change over time at measurable rates.

transmission would decrease. As early as 17 February 2020, Japan issued a guide with instructions on how to behave during the pandemic to avoid contagion and what to do upon the appearance of the first symptoms, discouraging early running of citizens supposedly infected to clinics, where people would be together in small spaces usually with poor ventilation and in close contact with others.

4. *Learning from experience and consideration of mental models influencing behaviour.*

I think that Western countries and Japan, or even Western countries and Asia, have fundamentally different ways of facing COVID-19, or even infectious diseases in general, including historical and cultural backgrounds. [I mentioned earlier that the] Western response was to identify cases and completely eliminate the virus. There is a notion of *completely annihilating the evil*. One way that is apparent is that not only politicians, but even many academic experts have used war metaphors to talk about COVID-19... I guess Japan and other Asian societies have developed a relationship with infectious diseases that contains a sort of resignation, as we had accepted living together with microbes (Hitoshi 2020).

Habit 6 of a Systems Thinker: Consider how mental models affect current reality and the future.

A Systems Thinker is aware of how beliefs and attitudes influence the way a system behaves.

Dr. Hitoshi says that attitudes, beliefs, ideas and perceptions are components of the structure of systems and affect their behaviour; that is, they affect both their perspectives and actions. The frame of mind described in the above quotation shows how the Japanese are coping with the pandemic, how they have historically

done so and how they will probably behave in the future, having learned from trends and patterns experienced with time.

Japan learned lessons from past outbreaks, especially from the 2009 H1N1 influenza pandemic. In 2009, people rushed to get tested, having to remain for hours in waiting areas characterized as closed spaces, crowded places and places that forced close contact (the three C's of the Japanese strategy). In 2020, the understanding that chaotic testing could make things worse led authorities to move even further away from identification of individual cases, instead focusing on clusters and massive spread.

A dramatic contrast with the West is worth mentioning. A recent book on the Spanish flu (Schwarcz and Starling 2020) highlights that difficulties and mistakes experienced during the new coronavirus pandemic could have been avoided had we considered lessons from what we experienced one century ago. Some examples are the paralysis of the economy; the fragility of public health systems; the extra burden on Black and poor people; the increase in social inequalities; the ineffectiveness of silver-bullet solutions such as hydroxychloroquine and ivermectin and the effectiveness of face masks, hand washing and social distancing.

By January and February 2020, and as time passed, strings of inter-related events unfolded in front of our eyes. Each of us witnessed events mutually affecting each other in various ways, in amazing feedback loops, dragging ourselves, our neighbours and our societies into uncertain and unpredictable situations. Think how many times you planned and postponed a trip or a vacation or asked yourself when you could go out with friends again. Think how your beliefs regarding viruses, infections, contagion, your rights, your feelings about the suppression of your daily freedom and the dynamics of your emotions in isolation from others affect how you experience the pandemic and the decisions you take. In short, *think of yourself as a system*, and grant to your structure the production of the behaviour you exhibit at each moment in time, per the flow of your life story in the contexts within which you interacted and the current inter-relationships with which you are engaged. Keep on reflecting and maintaining your stream of thought as close to you as possible, focusing on yourself, checking how interdependent you are with the multiple inter-relationships that you establish with systems in your context and they establish with you. Reflect on how the flow of your behavioural patterns is forming new patterns and trends and a whole new situation is emerging. It could not be otherwise. You and your circumstances are fully contingent phenomena.

'Which system are we talking about in this rather personal exercise?' you may ask. 'Whatever system you want to define as a unit to reflect upon at this moment,' we would reply. You determine the size and complexity of the system based on your interests, your capabilities and the problem or concern at stake.

Habit 7 of a Systems Thinker: Makes meaningful connections within and between systems.

A Systems Thinker sees how concepts, facts, and ideas link together, which can lead to new learning, discoveries, and innovations.

Habit 8 of a Systems Thinker: Considers short-term, long-term and unintended consequences of actions.

A Systems Thinker looks ahead and anticipates not only the immediate results of actions but also the effects down the road.

Habit 9 of a Systems Thinker: Recognizes that a system's structure generates its behaviour.

A Systems Thinker focuses on system structure and avoids blaming when things go wrong.

Habit 1 of a systems thinker proposes that we question how to balance the big picture with important details that may be raised. In other words, depending on how precise your question or problem is, the big picture can have more- or less-defined contours from the beginning, although none of the attitudes we have indicated here could be dispensed with, no matter how early you define the system you want to consider. Still, while reflecting, resist the urge to jump to conclusions. If you decide to deal with systems, you deal with complex units, and quick solutions are likely to fail. All phenomena that complex systems exhibit are complex and not immediately evident.

Now, how have scientists discussed SARS-CoV-2 and COVID-19? As of 29 January 2021, the World Health Organization (WHO) had officially reported 100,819,363 confirmed cases and 2,176,159 deaths, which indicates

Habit 10 of a Systems Thinker: Considers an issue fully and resists the urge to come to a quick conclusion.

A Systems Thinker takes the necessary time to understand the dynamics of a system before taking action.

a global pandemic of unprecedented magnitude². Accordingly, the new virus and the disease have received unusual attention from scientists; by the same date, there were 62,005 PubMed-indexed articles, with more than 6,000 publications in less than one month³. Since the virus was first reported in December 2019, 'the pace of investigation and publica-

tions makes SARS-CoV-2 the most-studied virus in history' (Baumgarth et al. 2020, 2342).

Baumgarth and colleagues (2020) indicate that communication of scientific work, especially on such a complex, extensive pandemic, is not trivial. Current communication has been accompanied by misunderstandings, leading to a lack of trust in science and to criticism that all biomedical areas of interest to the pandemic need to be reinvented. For example, scientists perceive that expressions of a nuanced approach to a particular question, such as the declaration that 'process X is poorly understood' or that 'there is a lack of detailed knowledge about something', indicate mature, high-integrity work, providing a sound rationale for studying the details in depth. On the contrary, the general public may perceive such statements as

² Available at <https://covid19.who.int/>.

³ Available at <https://tinyurl.com/3cz49p4e>.

indicating lack of expertise, which can lead to panic and denialist behaviour and can divert scientists from relevant discussions.

The pandemic has challenged epidemiologists, infectious disease specialists, immunologists, cardiologists, neurologists, intensive care personnel, medical doctors, biochemists and all. This pandemic offers the opportunity to acknowledge and to seek understanding from several fields to interpret a situation and to learn how to manage uncertainty and unpredictability. In addition, after 11 months, the approach to the disease at hospitals and clinics has changed, for much has been learned since its identification. Nevertheless, it is also correct to say that the context of the pandemic – with massive communication often resulting in quickly and widely spread and poorly interpreted information – has not helped populations follow rational precepts and protocols. Questionable results and alarmist news articles have filled the void generated by the inability of the scientific community to convey experimental results swiftly to the public, which is the proper way to discuss and advance science. This has hampered efforts to contain transmission, hindered development of therapeutics and furthered mistrust in vaccines, as Baumgarth and colleagues anticipated (2020, 2342–2343).

An example of rapidly spreading information in the beginning of the pandemic concerns the very nature of the disease. COVID-19 was initially – and continues to be – characterized as an acute respiratory problem. A few months later, when more cases were observed, pronounced lymphopenia was noted in severe cases – leading to an immediate characterization of COVID-19 as a hematologic (circulatory) disorder. A bit later, when cases indicating neurological disturbances appeared, the reference of the disease changed again. Later, it was summarized as a *systemic disease*. With time, refined, agreed-upon formulations appeared, but this is not a topic for us here. What is relevant in this moment is not the ‘scientific final word’ about COVID-19 but the development of a scientific understanding of the virus. Different characterizations of the disease were initially mutually exclusive, without paying attention to the simultaneous or secondary appearance of associated disorders. In this sense, the description of the disease as systemic addresses issues worth noting, such as the possibility that systems thinking gives us to approach a phenomenon in its multiple aspects and from multiple perspectives at the same time, even if

Habit 11 of a Systems Thinker: Surfaces and tests assumptions.

A Systems Thinker actively tests theories and surfaces assumptions, perhaps with others, in order to improve performance.

they seem paradoxical at first. Moreover, it helps us acknowledge that the variety of symptoms and configurations of this disease in different people depends on the structure of the organism when the infection occurs and its functioning through time, together with the viral load.

**Habit 12 of a Systems Thinker:
Changes perspectives to increase
understanding.**

A Systems Thinker increases understanding by changing the way they view aspects of the system.

From the perspective of global systems, such as the global economy, and prevalent neoliberal politics that broadly determine contemporary life dynamics, how can we perceive the development of the pandemic? Criticisms of neoliberal politics, market-driven economies and societies focused on individualism increased in intensity when

the impact of the pandemic became clearer. On 3 April 2020, the *Financial Times* (2020) acknowledged that the coronavirus exposed frailties in current economic and social models, and the Editorial Board advocated for reforms:

Radical reforms – reversing the prevailing policy direction of the last four decades – will need to be put on the table. *Governments will have to accept a more active role in the economy.* They must see public services as investments rather than liabilities and look for ways to make labour markets less insecure. *Redistribution will again be on the agenda; the privileges of the elderly and wealthy in question. Policies until recently considered eccentric, such as basic income and wealth taxes, will have to be in the mix* [emphasis added].

The proposals are surprising, considering that the *Financial Times* is a strong voice of the neoliberals:

If there is a silver lining to the Covid-19 pandemic, it is that *it has injected a sense of togetherness* into polarised societies. But the virus, and the economic lockdowns needed to combat it, also shine a glaring light on *existing inequalities – and even create new ones.* Beyond defeating the disease, the great test all countries will soon face is *whether current feelings of common purpose will shape society after the crisis* [emphasis added].

The call for strengthening a sense of togetherness and the observation that something new in this direction may have been injected during the pandemic recalls Schwarcz and Starling's (2020) narrative about how attitudes of Brazilians during the Spanish flu contrasted with their attitudes

during the current pandemic. At the time of the Spanish flu, after a period of lockdown, the owner of many movie theatres in Belo Horizonte, one of the Brazilian state capitals, decided to reopen, claiming that radical protection measures were unnecessary and highlighting his financial losses. He promoted the first show after the lockdown and argued that the spirits of the population needed to be raised in times of a pandemic. What he did not foresee was the strong reaction of the population, who proposed boycotting the shows; as a result, the movie theatres had no audience until the influenza was gone. In contrast, in 2020/21, a common scenario in Brazil, France, Spain, the Netherlands, Denmark and many other countries is protests against social distancing and in favour of market opening and the right to go to crowded parties, bars and beaches. What the *Financial Times* editors suggest is happening, therefore, seems not to be the case. We could indicate the influence of mental models in this behaviour as well – a more individualistic mental model with a reduced sense of togetherness seems now to be dominant in many parts of the world.

The pandemic has exposed inequalities of all sorts. According to Human Rights Watch (2020), in the United States, the outbreak has highlighted economic inequalities and the fragile social safety net that leaves vulnerable communities to bear the economic brunt of the crisis. Although the virus infects people from all social classes, the poor are the most affected because of long-standing segregation according to income and race, limited economic mobility, poor facilities and the high cost of medical care. These observations can be extended around the planet. In countries like Brazil, the disease is prevalent among Blacks, Indigenous people, the poor and those living in crowded spaces and in areas without basic sanitation. The opening of markets placed an extra burden on those who depend on overcrowded public transportation to travel from home to work and back, expanding the chances of contagion. A cruel inequality derives from the digital segregation of Black and poor people observed in Brazil and other countries. When classes occur primarily virtually, a significant number of children and young adults enrolled in school have remained without schooling for the whole period (Mari 2020).

Many other global challenges related to the SDGs, such as unemployment; hunger; disruption of formal education; growth in domestic violence;

Habit 13 of a Systems Thinker: Recognizes the impact of time delays when exploring cause and effect relationships.

A Systems Thinker understands that often cause and effect are not closely related in time.

emergence or deepening of psychological disturbances and collapse of public health systems, the economy and the finance systems, have arisen. These dramatic occurrences can be systemically understood as the behavioural paths these systems are taking, after changes in their context of interactions and in their structures.

Habit 14 of a Systems Thinker: Uses understanding of system structure to identify possible leverage actions.

A Systems Thinker uses system understanding to determine what small actions will most likely produce desirable results.

Understanding the structural aspects of this pandemic identifies key issues that provide a dire warning for the future emergence and spread of zoonotic diseases (that transfer from animals to humans). This includes the historical, grave exploitation of the environment; biodiversity loss; deforestation; dietary habits and agricultural defences. The con-

nections between these structural aspects cannot be disregarded or set aside as unimportant. On the positive side, the experience has provided an opportunity for new indicators to support the SDGs and a green restoration (UNEP 2020). Navaratnam-Blais (2020) highlights a potentially fruitful convergence of two agendas that has emerged in organizations – fight climate change and deepen digitization:

These two agendas will exist in something of a symbiotic relationship; digitization will allow companies to meet their decarbonization targets, while the pressures of climate change will help create the business case for accelerated investment into digital transformation. In October of this year, for example, we at Source Global Research published a report, based on a survey of 150 senior US executives, exploring how various organizations intend to use professional services firms to help them mitigate their exposure to climate risk and achieve their decarbonization targets. When those executives were asked what specific steps their businesses could take that would make the greatest contribution towards those objectives, the most popular answer – by quite some margin – was ‘finding innovative ways to incorporate green technology into our operations’.

In systems thinking terms, if his observations and hopes are confirmed, the tech-celeration that is taking place during the pandemic may be producing further effects and establishing new interrelationships between systems that were not before linked.

We are all undoubtedly capable of enlarging and unfolding this complex, interrelated network, contributing with personal, contextual and diversified perspectives on the enormous fabric that constitutes our life experience. 2020 – and it seems that the same will continue during 2021 and 2022 – has whipped the planet overwhelmingly, with contingencies that depend on individual characteristics; social, cultural and economic backgrounds; politics and history. All in all, 2020 has provided an extraordinary experience, at the same time collective and individual, in which it is clear that we are learning how to live with uncertainty and unpredictability. Although the experience of lack of control and predictability has often been painful, it demands our immediate attention, inspiring us to asking appropriate questions and reformulating our common way of reasoning.

Quick Takeaways

With what we know about SARS-CoV-2 and the COVID-19 pandemic, and with adoption of a systems thinking perspective, numerous practical learning examples and approaches for ongoing concerns, at all levels, could be immediately pointed out:

1. Experience has shown that, when lockdown restrictions are relaxed because of a decrease in transmission, an increase in new cases follows. The recurrence and escalation of this vicious circle threatens economies and societies and must be stopped. This phenomenon can be referred to as the circular nature of complex cause-and-effect relationships. Time delays may affect them, requiring that we acknowledge that, in systems, cause and effect may not be closely related in time and not even in space. This sort of phenomenon can also be treated as feedback loops, shedding light on recursiveness. Recursiveness appears when we seriously consider the flow of time – or history. It means that each new action taken, for example a lockdown, will fall on a market already weakened by a previous lockdown, deepening further economic frailties. New actions do not fall on the initial state of the system.
2. Systematic consideration of the dynamics of systems and interrelationships with the context alerts us that the focus on flows of interrelationships as historical processes is crucial. In the case of the pandemic, the focus on high contagion potential yields to a variety of possible solutions that will work only if local habits, culture and beliefs are taken into consideration.

- Regarding recovery of urban circulation and reopening of markets, places and attitudes that require or involve close contacts, crowded places and closed spaces need to be avoided. Architectural solutions allowing for cross-ventilation could be found for originally closed spaces such as temples and classrooms, for example.
 - Regarding vaccination campaigns, setting priorities requires the consideration of various interrelated factors, that are highly contextual and dynamic. The process of establishing the target clusters themselves involves a close examination of the dynamics of the society, a clear understanding of the concerns of national authorities and the population, an assessment of the resources available and an overview of demographics. One size does not fit all. Examples from various countries can guide, but not determine, what is appropriate for different systems' structures and dynamics. Throughout the world, governments have exhibited different concerns, which then result in different priorities setting as, for example:
 - *to stop the flow of the contagion*: in this case, those who must go out to work every day could be considered immediate priorities. This was done in Indonesia;
 - *to reduce deaths*: in the beginning of the pandemic deaths occurred mainly among elderly people, and because vaccines were scarce, some countries opted to immunize the elderly living in nursing homes. What was not considered was that reinfections could occur and caretakers would come and go every day, keeping the potential for infection the same. Moreover, in unequal countries like Brazil, Blacks, Indigenous people and the poor, who lack the resources to pay for a nursing home, who have lower life expectancies than whites, were left unattended although they are among the most vulnerable groups which were, then, unattended;
 - *to immunize the most vulnerable groups first*.
3. Outreach campaigns to raise awareness of the effects of an individual's attitudes on the community as a whole, and vice versa, are paramount, because a chain of interrelated events occurs as a consequence of each individual's performance. Following safe practices can help avoid the collapse of health systems and institutions, keeping hospital facilities and intensive care units available to treat extreme cases, reducing deaths and burden on the health

care system and enabling the opening of markets and flexibilization of social interaction in the medium term. This awareness can build collectively through beneficial virtuous circles.

4. Given that COVID-19 may have long-term effects, individuals who were infected may see their productivity decrease after discharge, and their need for specialized care increase, which burdens the family, the country economy, health care system and social well-being. Population immunity (also called herd immunity) requires that 70 per cent to 80 per cent of the population be infected or vaccinated. Therefore, controlling the pandemic by letting the population become infected is not a good option, considering the possible medium- and long-term consequences for individuals, their families and the country. According to the WHO, the United States, the country with the most deaths and identified cases by January 2021, has detected 25,354,044 cases⁴ in a population of 330,053,524 (U.S. Census Bureau 2021). This amounts to only 7.6 per cent of the population. Before the distribution of vaccines, the country experienced an extremely serious situation, having to deal with a percentage of infected people which is much lower than what is required for population immunity. Although it is a huge challenge to produce and distribute vaccines for 70 per cent to 80 per cent of the population, stimulating the spread of the virus would mean that the country would need to undergo many more waves of infection to achieve herd immunity, with unforeseen short, medium and long-term consequences.
5. Technological solutions have provided significant support during this first experience with a pandemic of this magnitude, to the point that this process was coined 'tech-celeration' (Standage 2020). Scientists such as Dr. Hitoshi (2020) alert us that additional surges are likely to come, perhaps involving new viruses. Incentives and clear policies to expand access to the Internet, to produce and finance digital devices for reasonable prices to cover more populations, especially the poorest and most vulnerable, can be designed and implemented. The inequalities that COVID-19 has exposed and exacerbated may be dramatically amplified in the medium to long term, leading to further economic and social collapse, with

⁴ Updated map available at <https://covid19.who.int/table?tableChartType=heat>.

consequences for the planet itself if immediate attention is not paid to contain this process.

6. The urgent call to transform the relationship between humans and their environment, living and non-living, has become even more urgent during this pandemic. The disruption of the equilibrium of the planet during the Anthropocene era has as one of its consequences the appearance of strong, new viruses of animal origin with high infection potential. Evaluators can strongly support goals and interventions aimed at reversing the imbalance of the Anthropocene and promote the transformations required.
7. Finally, as Patton (2020b) recommends, systems thinking should be systematically embraced in evaluation, and we should become real fact checkers.

11. Model systematic evaluative thinking. The media are filled to overflowing with opinions about what's working and not working, what's been done well and poorly, and who's to blame and who gets credit. Everyone is an evaluator. But we are professional, systematic evaluators. Evaluate for yourself – with skill, care, and thoughtfulness – what's working and not working to mitigate the crisis. Be prepared to render judgments as appropriate based on cumulative evidence, but also be prepared to demonstrate evaluative thinking when evidence is inadequate, when judgments are premature, and when the facts are uncertain. Refrain from expressing uninformed or premature judgments and urge others to do likewise.

Transforming Evaluation with a Paradigmatic Change

Exploring the habits of a systems thinker in light of the COVID-19 pandemic can reveal numerous themes that are of concern in development and evaluation, such as resilience, adaptiveness, sustainability and transformation. It can offer alternatives for transforming evaluation, reframing its object and the evaluation criteria and revealing a way out of dilemmas that evaluators face, especially considering the commitment to achieving the SDGs (Ofir et al. 2019).

Patton's blogs (2020b; 2021) make recommendations to evaluators facing the current global health emergency that can be expanded to evaluators' work beyond the pandemic crisis. His recommendations are made from a systems perspective and can be identified through the journey

with the systems thinkers' habits. As we list takeaways from the journey, it becomes clear that, to be fluent in systems thinking, it is necessary to uphold the full potential of systemic evaluation approaches (see e.g. Patton 2020c) and that the Principles for Effective Use of Systems Thinking in Evaluation (SETIG 2018), which are in line with the reasoning expressed through our journey, are useful in a dynamic, interrelated, indivisible way.

To take the system and the context into account often seems like a huge task that involves complications besides complexity, is impossible to complete and is arbitrary and biased in its reach and understanding and hopelessly amateurish. It is true that uncertainties exist and that, in some cases, data are unreliable, unverifiable and not much better than guestimates, but this is often the case at the intervention level as well. Interventions aimed at transformational change need to have information and data on and insight into the systems that are targeted; otherwise, they are set up for failure.

An important issue is the use of criteria to decide whether an intervention managed to set transformative change in motion and, if so, whether that change was for the better rather than making things worse. It is clear that the Development Assistance Committee (DAC) criteria were not formulated for interventions in the context of systems change, let alone transformational change. Patton (2020c) indicates how a systems imperative would translate into new criteria to help make evaluations relevant for transformational change while pointing out that the traditional DAC criteria, although redesigned, support business as usual and do not meet current global needs. There is a lively, ongoing debate on the matter worth joining that focuses not only on the global crises of our times, but also on social justice, human rights, exploitation, conflict and violence.

As evaluators, we are used to starting an evaluation by providing validated data regarding a country, a region or a locality, including historical perspectives. What we need to do for transformational change is to provide similar data and insight for the system that is supposed to be changed, whether it is a market system, the interaction between human activity and an ecosystem or any other well-identified system that needs to transform. National bureaus of statistics, as well as many global databases of international organizations and a plethora of research and science programmes at universities around the world, contribute to our understanding of key systems and their contexts but do not substitute for clear systems reasoning. We advocate that the habits of a systems thinker be used to explore these sources to start an exploration of transformations that societies should strive for. This should become an integral part of the new paradigm of a transformed evaluation for transformation.

Note

This chapter was finalized when Humberto Maturana passed away on 6 May 2021. For all those seeking innovative ways to understand our present and build a better future we recommend his ideas.

Acknowledgements

Our heartfelt thanks to the anthropologist and linguist Victor Paredes-Castro; Dr. Nelson Vaz, immunologist and emeritus professor at Federal University of Minas Gerais, Brazil; and Sheri Marlin, Chief Learning Officer at the Waters Center for Systems Thinking and co-author of *The Habit-Forming Guide to Become a Systems Thinker* (Benson and Marlin 2019), for the enlightening suggestions, references and comments.

References

- Baumgarth, Nicole, Janko Nikolich-Zugich, F. Eun-Hyung Lee et al. 2020. Antibody responses to SARS-CoV-2: Let's stick to known knowns. *Journal of Immunology* 205: 2342–2350.
- Benson, Tracy, and Sheri Marlin. 2019. *The habit-forming guide to become a systems thinker*, 2nd ed. Pittsburgh: Waters Center for Systems Thinking.
- Financial Times. 2020. Virus lays bare the frailty of the social contract. Radical reforms are required to forge a society that will work for all.
- Hitoshi, Oshitani. 2020. Infectious disease response – to see the forest, not just the trees: What differentiated Japan from the Western countries? *Diplomacy* 58, 5 June.
- Human Rights Watch. 2020. US: Address impact of Covid-19 on poor. Virus outbreak highlights structural inequalities. <https://tinyurl.com/yj8k5r4x>
- International Science Council. 2017. Report: A guide to SDG interactions: From science to implementation. <https://tinyurl.com/wlyx7yw>
- Latour, Bruno, and Steve Woolgar. 1986. *Laboratory life: The construction of scientific facts*, 2nd ed. Princeton: Princeton University Press.
- Latour, Bruno. 1987. *Science in action: How to follow scientists and engineers through society*. Los Angeles: Sage.
- Latour, Bruno. 1988. *The pasteurization of France*. Cambridge, MA: Harvard University Press.
- Lewin, K. 1952. *Field theory in social science: Selected theoretical papers by Kurt Lewin*. London: Tavistock.

- Magro, Cristina, and Rob D. van den Berg. 2019. Systems evaluations for transformational change: Challenges and opportunities. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob van den Berg, Cristina Magro and Silvia Salinas Mulder, 131–152. Exeter, UK: IDEAS.
- Mari, Angelica. 2020. Negros e pobres sofrem com exclusão digital durante a pandemia. *Revista Forbes* 27 Maio. <https://tinyurl.com/yy2q96by>
- Maturana, Humberto, and Francisco Varela. 1973. Autopoiesis: The organization of the living. In: *Autopoiesis and cognition – the realization of the living. Boston studies in the philosophy of science*, vol. 42, (eds.) Humberto Maturana and Francisco Varela. Dordrecht: Reidel Publishing.
- Maturana, Humberto and Francisco Varela. 1992. *The tree of knowledge: The biological roots of human understanding*. Boston and London: Shambhala.
- Navaratnam-Blair, Fergus. 2020. Sustainable transformation: Why climate change and the digital agenda are more closely linked than ever. *Information Age* 12 November. <https://tinyurl.com/y2r8r6oq>
- Ofir, Zenda, Gerald Singh, Emilie Beauchamp et al. 2019. *From monitoring goals to systems-informed evaluation. Insights from SDG14*. IIED Briefing Paper 11. <https://tinyurl.com/yy3gmmbe>
- Patton, Michael Quinn. 2020a. *Blue Marble Evaluation: Premises and principles*. New York: Guilford Press.
- Patton, Michael Quinn. 2020b. Evaluation implications of the coronavirus global health pandemic emergency. <https://tinyurl.com/t6th337>
- Patton, Michael Quinn. 2020c. Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation* 39 (4): 527–542. <https://doi.org/10.1177/1098214020933689>
- Patton, Michael Quinn. 2021. Lessons about evaluating responses to the pandemic: Insights from the World Food Programme Developmental Evaluation. <https://tinyurl.com/yk5scrvq>
- Rowe, Andy. 2012. Evaluation of natural resource interventions. *American Journal of Evaluation* 33: 384. <https://doi.org/10.1177/1098214012440026>
- Roy, Arundhati. 2020. The pandemic is a portal. *Financial Times* 3 April. <https://tinyurl.com/saed22p>
- Sander, Elizabeth, and Oliver Bauman. 2020. Five reasons why Zoom meetings are so exhausting. *The Conversation*. <https://tinyurl.com/y65klnuf>
- Schwarzc, Lilia Moritz, and Heloisa Murgel Starling. 2020. *A bailarina da morte: A gripe espanhola no Brasil*. Rio de Janeiro: Companhia das Letras.
- SETIG (Systems in Evaluation Topical Interest Group of the American Evaluation Association). 2018. Principles for effective use of systems thinking in evaluation practice.
- Standage, Tom. 2020. Ten trends to watch in the coming year. A letter from Tom Standage, editor of *The World in 2021*. <https://tinyurl.com/y6jsjug2>
- Uitto, Juha I. (ed.). 2014. *Evaluating environment in international development*. London: Routledge. <https://doi.org/10.4324/9781315814025>

- Uitto, Juha I., Jyotsna Puri, Rob D. van den Berg (eds.). 2017. *Evaluating climate change action for sustainable development*. Springer Open. <https://tinyurl.com/y5eh3e3c>
- UNEP (United Nations Environment Programme). 2020. COVID-19 response. <https://wesr.unep.org/>
- United States Census Bureau. 2021. U.S. and World Population Clock. <https://www.census.gov/popclock/>
- Waters Center for Systems Thinking. 2019. Habits of a systems thinker. <https://tinyurl.com/3yw5vz8b>
- Waters Center for Systems Thinking. 2021. What is systems thinking? <https://tinyurl.com/3mw4mhsn>

PART VI

THE
PRAGUE
DECLARATION





CHAPTER 18

The Prague Declaration: Meaning and Testimonials

ROB D. VAN DEN BERG, DANIEL SVOBODA, ADA OCAMPO,
JUHA I. UITTO, SILVIA SALINAS MULDER, RASHMI AGRAWAL
AND JOSEPHINE WATERA

The International Development Evaluation Association (IDEAS) has a long history of discussing global and international issues in development. From 2015 onwards this focused on how evaluators should take up sustainability issues in their work (Bangkok, October 2015), to how evaluators could support progress towards the Sustainable Development Goals (Guanajuato, December 2017), to how evaluation could contribute to transformational change to solve the global crises of our times (Prague, October 2019). In Prague, the IDEAS Global Assembly was joined with the Third International Conference on Evaluating Environment and Development, organized by the Community of Practice EarthEval and the Independent Evaluation Office of the Global Environment Facility. This led to a 'perfect storm' of ideas on how evaluation could support and strengthen transformational change, from economic, social to environmental issues, taking into account equity and equality as well as working in contexts of fragility, conflict and violence. Many voices from the Global South were welcomed in Prague and special sessions were held to discuss local issues with global consequences and global problems impacting on local conditions. The time was ripe to harvest insights, connections and opportunities.

The Prague Director of the two conferences, Daniel Svoboda, proposed at an early stage to think about the possibility of adopting a 'Prague Declaration' at the end of the meeting. A first discussion of this idea took place at a meeting of the Czech Evaluation Society. Furthermore, it was taken up as a pre-conference workshop, where brainstorming could take place to discuss the possible content of the declaration. A special session was also planned during the conference to fine-tune the draft declaration, attended by a large group of interested evaluators. The resulting text was presented at the closing session of the conference – read by Rob van den Berg and Juha Uitto to all present participants – and accepted by acclamation.

This chapter first presents the Prague Declaration. It then gives the floor to Daniel Svoboda for his personal perception of the Declaration and what it meant for him and for countries in transition in Central and Eastern Europe. After this, three short testimonials of colleagues are presented to add perspectives. The last section of this chapter brings short reflections of the current President of IDEAS, Ada Ocampo, of Juha I. Uitto, Director of the Independent Evaluation Office of the GEF and of Rob D. van den Berg, President of IDEAS at the time of the conference.

Prague Declaration on Evaluation for Transformational Change

ADOPTED ON FRIDAY 4 OCTOBER 2019

We, the evaluators, commissioners, parliamentarians and other evaluation users, gathered in the IDEAS Global Assembly and the Third International Conference on Evaluating Environment and Development, recognize the need and urgency of systemic change from local to global levels to address the global crises endangering our future. Having discussed the role of evaluation in promoting learning, systemic and transformational change, we agree on the following statements.

1. Promote Transformational Evaluation for the Sustainable Development Goals

We commit to evaluations that help us learn, understand and support the transformational and systemic changes needed in our countries and the world, as agreed upon in the 2030 Agenda for Sustainable Development. A sustainable balance between the social, economic and environmental domains is crucial in light of the existential threats of the climate crisis, mass extinction of species, growing local and global inequity, and ultimately unsustainable use of the resources of the planet.

2. Work in partnership

We will promote partnerships among evaluators, based on applied ethic codes and professional standards, and on mutual trust.

At the same time, we commit to engage and recognize new evaluators and collaborators from many different disciplines and fields of work, including young and emerging evaluators, students and interns in evaluation teams whenever possible, in order to promote mutual learning and to discover and leverage new views and skills.

3. Explore power relations and promote inclusiveness

We will deal sensitively and effectively with the unequal power relations that are apparent throughout intervention and evaluation processes. We

commit to applying approaches that include the marginalized, and to respecting the need to engage local stakeholders in consultations about the purpose of evaluations, evaluation questions, and preliminary conclusions and recommendations. Looking towards a connected future of evaluation, we commit to co-designing and co-conducting evaluations that include indigenous and local ways of knowing with conventional and transformational methodologies.

4. Respect for rights and responsibilities

In all our approaches, communications and deliverables, we commit to respecting privacy, equity, gender equality, minorities and indigenous peoples, the dignity of people and environmental integrity. We commit to respecting and advancing human rights and responsibilities, as well as the rights of societies and of nature.

5. Support for professionalization and capacity development

We advocate a transformational change of evaluation itself. We will support efforts to bring knowledge and capacities to commissioners, evaluators, development partners and the diversity of stakeholders who can and do contribute to the practice of development evaluation throughout the world.

We support the development of an international evaluation academy to advance professionalization and promote the interaction between science, research and evaluation to enrich our profession and our efforts to support evaluation capacity at all levels.

6. Focus on sustainability

In all our evaluations, we commit to evaluating for social, environmental and economic sustainability and transformation, including by assessing contextual factors and systemic changes. We commit to assessing and highlighting, in all evaluations, unintended negative social, economic and environmental effects.

7. Focus on fragility, conflict and violence (FCV)

We commit to understand and work on the dividers and connectors of conflict and violence and apply evaluation approaches that are gender and conflict-sensitive and based on the principles of 'Do No Harm' as described in the IDEAS Guide on Evaluation in Fragility, Conflict and Violence, as discussed during the Global Assembly.

8. Support for transformational indigenous¹ evaluation

We commit to value and support the strengthening of and learning from indigenous evaluation by and for indigenous peoples.

9. Shared responsibility for results

We fully understand that the real result of an evaluation is not the evaluation itself but the use that is made of the evaluation in all of its phases. We commit where possible to work with the potential evaluation users, including on possible solutions for problems identified.

10. The challenge ahead

The discourse on evaluation for transformational change is challenging for evaluators working in systems, contexts and circumstances that are not yet open to or sufficiently enabled to commit to transformational evaluation, and challenging for commissioners, users and stakeholders. We continue to discuss and to deepen our understanding of the changes required for evaluation to contribute to tackling the crucial problems of our time.

¹ After the Prague conferences, the use of a capital for the word Indigenous became standard. While we have not changed the text of the Prague Declaration, we use the capitalized Indigenous elsewhere in this volume.

The Need for Partnership in Using Evaluative Evidence for Transition

DANIEL SVOBODA

Introduction

Partnership is an important word, widely used for decades, even centuries. Partnership in development is being redefined and gaining importance. Since the Paris Declaration on Aid Effectiveness was adopted in 2005 (OECD DAC n.d.), we have been officially using the term ‘partner countries’ instead of ‘aid recipients’. In 2008, the signatories of the Accra Agenda for Action confirmed their commitment ‘to eradicating poverty and promoting peace and prosperity by building stronger, more effective partnerships that enable developing countries to realise their development goals’ (OECD DAC n.d.). In 2012, all key development actors reached a consensus on the Global Partnership for Effective Development Cooperation (GPEDC 2012), a multi-stakeholder platform that brings together all types of development actors to increase the effectiveness of their development efforts, deliver long-lasting results and contribute to achievement of the Sustainable Development Goals (SDGs).

The United Nations (2005, 4) defines partnerships as ‘voluntary and collaborative relationships between various parties, both State and non-State, in which all participants agree to work together to achieve a common purpose or undertake a specific task and to share risks and responsibilities, resources and benefits’.

In the field of development evaluations, many evaluators have established strong partnerships with evaluation users and other development actors. Many evaluation users are cooperating closely with evaluators because they see the importance of good evaluations.

Partnership of diverse actors is the key precondition for any transformational change, as well as for reaching the SDGs. According to SDG 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development:

A successful sustainable development agenda requires partnerships between governments, the private sector and civil society. These inclusive partnerships built upon principles and values, a shared vision, and shared goals that place people and the planet at the centre, are needed at the global, regional, national and local level (United Nations 2015).

How virtual or how real are these proclaimed partnerships? Who are the partners? What are the common goals, shared responsibilities or hidden interests? What are the key principles of partnership? How can the partnerships help in using evaluations?

In this section of the Prague Declaration chapter, I propose possible responses and a few more questions. I would like to open a discussion about several important principles of partnerships, the diverse partners and special aspects of partnerships in evaluations. For illustration, several real-life stories from my professional career are added to each point.

We widely use some additional important terms: *transition, transformation, behaviour change*. What do these words really mean, and what do they have in common? How do they relate to partnership? Any sustainable change depends on peoples' behaviour. Even the most expensive technological solutions or perfectly substantiated evaluation results will not work effectively if people will not use them.

Evaluation is a potent tool for identifying the motivations, critical assumptions and barriers, and systemic challenges or imbalanced power relationships that can complicate any transformational change. Partnership in evaluation is a feasible way to overcome these challenges by using watertight evidence, informed analysis of contextual factors and the empowering role of evaluations together. Such partnerships call for mutual accountability and shared responsibility but also bring many mutual benefits, including from mutual learning.

In totalitarian regimes, any change can be enforced, regulated and controlled without evaluation. People and their families can be punished (and sometimes even killed) if they do not follow the rules, and supporters and informers can be rewarded (usually only in the short term, until they become enemies of the regime). Can such forced behaviour produce positive, sustainable development change? This is impossible for many reasons, among them:

- Such approaches abuse all fundamental human rights and freedoms; the people are not at the centre, and many actors are intentionally left behind.
- There is no ownership of the change – the objectives have not been agreed upon in a participatory way, and thus the people will not (and usually cannot) participate in the development process either.

- Neither the analyses of problems nor monitoring and evaluation can be impartial, and thus the decisions cannot be evidence based.
- No one is responsible for the results or sustaining the benefits.

Accordingly, the expert and facilitation roles of evaluations are crucially important in fragile and transitioning countries, where support for transformational change is most needed.

Case 1: Speaking from the Czech experience, our development evaluations started in 2003, jointly with the effort to transform the Czech Official Development Assistance system, fragmented at that time under 11 sectoral ministries. The Ministry of Foreign Affairs commissioned these first evaluations, but the line ministries as the anticipated evaluation users were not engaged in preparations of these evaluations and not always interested in their results. At that time, the United Nations Development Programme Regional Bureau for Central Europe and Commonwealth of Independent States managed the evaluations.

This methodological (and financial) support enabled creation of a small evaluation unit at the Ministry of Foreign Affairs a few years later. It then took several more years to adjust the evaluation system, including establishing the Working Group on Evaluations within the Council for International Development Cooperation, creating the Reference Group on Evaluations, introducing new templates, formalizing the system for consultations between stakeholders to finalize the evaluation results, publishing the evaluation reports on the website and presenting the recommendation tracking system to all key stakeholders.

All these achievements would have been much more difficult without multi-stakeholder dialogue and international support and exchange of experience.

Significant changes are also necessary for sustainable development at the global level – to protect our Blue Marble Planet and mankind. The unprecedented COVID-19 pandemic that has affected all of us has highlighted the interconnectedness of our lives and thus the need for global solidarity and global responsibility. On the other hand, the pandemic has also brought a unique opportunity for reconsidering our priorities and introducing new ways of cooperation.

Most of the key aspects of partnerships in evaluations are reflected in the Prague Declaration on Evaluation for Transformational Change (IDEAS 2019). The consensus that diverse development actors reached at the 2019 IDEAS Global Assembly is a promising start on the way forward.

Key Principles of Partnership in Evaluations

Nobody can simply decide what is right and what is wrong, and nobody can succeed alone. Working together is the only chance for any transformational or systemic change. We must work in partnership with like-minded people, we must learn from each other and we must be open to communication and cooperation with all development actors.

All partnerships have some rules and principles, and there are many definitions of core partnership principles. These are usually based on aligning the interests of partners around a common vision (*convergence of interests and motivations*), combining their complementary resources and competencies (*complementarity of resources and approaches*), sharing accountability and risk (*mutual accountability*), maximizing value creation to achieve common goals and delivering benefits to all partners (*shared values*) (see e.g. The Partnering Initiative and UN DESA 2020; UN University Institute for the Advanced Study of Sustainability 2018; World Bank 2004).

The above principles respond to the questions WHY (common motivations to reach the foreseen goals and shared values) and HOW (everybody must contribute and be accountable). All these principles are necessary and can be confirmed in written partnership agreements, but they are not sufficient. I would like to highlight some 'soft' principles that are necessary at the operational level but cannot be achieved merely by signing an agreement. I also propose some (quasi) indicators (in *italics*) of whether these principles are working. The importance of these aspects is documented using several real-life examples and by my personal opinion (in boxes).

Mutual Trust

The first soft principle of partnership is *mutual trust*. It usually takes years to create trust, and trust can disappear in a minute, after a single mistake. Without trust, cooperation needs to be enforced, and the right to take the initiative is missing. Lack of trust undermines the effectiveness of the whole process and upsets the balance of the power.

For measuring trust, I propose the following quasi-indicator: The partners are not afraid to confess their own mistakes

It is definitely good if the partners share a vision, but there is a question: 'Whose vision is it really?' It is probably better if the partners keep and adhere to their own visions (if these are not contradictory) and can still trust each other and work together to reach common goals. The visions are usually anchored in a historical heritage of habits, fears and dreams, as well as in actual context (economic, social, environmental, political). Such heritage can hardly be shared. '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?' (Aronsson and Hassnain 2019, 92).

Case 2: We have been working in Vietnam for many years, and our cooperation was quite successful when considering the official indicators, but the clearest evidence of trust was a moment when our local partners admitted that both sides made some mistakes. It took two years from project implementation. Only after that were we able to confirm the common goals and to agree on the most effective ways to reach them.

If we cannot trust our partners, we must do most of the work ourselves and cannot exploit the benefits of partnership.

Fairness

Then it would be easy to use the proposed indicator: The partners can rely on each other.

A closely related principle is *fairness*. The partners must have similar ethics and code of conduct (not necessarily the formal one, hanging on the wall), avoid biases and be honest in communicating their own expertise and limitations. There must also be sufficient transparency at all steps during the evaluations (see e.g. IDEAS 2014).

Case 3: I remember a junior evaluator who participated in my training. A week later, she submitted her bid to a tender where she mentioned that she had organized that training.

Dishonest persons cannot be fair evaluators.

Shared Responsibility for Results

Another principle, complementing mutual accountability, is *shared responsibility for results* (foreseen and unintended impacts of each evaluation). The partners must work together with the same aim, and they must trust and defend their joint results. Those who will apply them should be consulted in the development of recommendations of any evaluation to ensure their applicability. The evaluation users should be included in the partnership schemes.

Proposed basic indicator: The partners are aware of the use of their results.

Opinion 1: One of the easiest steps is publishing the evaluation results (without classified information). Then everybody can easily assess their usefulness and monitor the application of evaluation recommendations.

Evaluators should follow how their results have been used.

Professional Standards and Using Lessons Learned

The partners must apply *professional standards* and be interested in *using lessons learned* for continuous improvement. This is usually an integral part of any professional code of conduct (see e.g. IDEAS 2012).

There must be a learning mechanism in place. It can take diverse forms: internal quality assurance or formal checklists on the quality of the evaluation reports, consulting a reference group during the evaluations, voluntary peer reviews, any kind of accreditation. In my opinion, voluntary peer reviews and consequent use of lessons learned is the best way to learn.

Proposed quasi-indicator: Partners are ready to correct their own mistakes together.

Case 4: I worked for many years in the Reference Group on Evaluations of the Czech Ministry of Foreign Affairs. Many young and emerging evaluators were grateful for the feedback they received, and they improved substantially afterward, but many renowned evaluators (with certified quality management systems) repeated the same mistakes and held the same biases every year. This can be also seen in the mandatory responses to the comments from the Reference Group in the published evaluation reports.

Any expertise needs a continuing learning.

Inclusiveness

Finally, the evaluation partnerships must be *inclusive*. No one can be left behind, and no critical assumption or contextual factor can be omitted. These factors cannot be properly identified without engagement of local actors and evaluation users.

Proposed indicator:
All key stakeholders
have been engaged
in the evaluation
process.

Directly engaged groups in an intervention have the greatest knowledge of the overall context. They can recognize the real successes and failures and identify sustainability issues. The evaluation users should be aware of the needs of, positions of and constraints on other groups, and likewise, the evaluators should be aware of their clients' and evaluation users' situation. Working together from the beginning of the evaluation process can increase the impact of the evaluation.

Case 5: One of my first international evaluations was on a large, 10-year-long project in Palestine. The ministry responsible for this project (the main addressee of the evaluation recommendations) did not wait for the return of the evaluation team from the field mission and launched the call for continuation of the project without any reflections on the evaluation findings, conclusions and recommendations. One of the main reasons for ignoring the evaluation results was that this ministry was not sufficiently engaged in preparation of the evaluation design and even refused to participate in the field mission.

The principle of inclusiveness concerns not only the target groups of an intervention, but also the final users of evaluation results.

Partners in Evaluations

Who are the partners in evaluations? There are many levels of partnerships, but not all have been sufficiently supported.

Partnership within the Core Evaluation Team

Do all members of the team know the real purpose of their evaluation? Do they trust each other? Do they back up and defend each other? Do they discuss their mistakes and correct them?

Evaluation is teamwork. It is difficult to work in a team if there is no previous experience of cooperation, no trust established and no rules regarding responsibility established. Some commissioners select evaluation team leaders and team members independently, according to their biographies or financial offers. This is a lottery that might be useful for a sampling strategy but will not contribute to reliable evaluation results.

Evaluation cannot have useful results if the chemistry within the team does not work and the people do not know and trust each other.

Case 6: A few years ago, I was working on an international evaluation team of which all members were selected independently. Unfortunately, the team leader did not follow his duties and deadlines, and even more, he distorted the findings from his own surveys (replacing the real responses with his own opinion). This complicated the work of the team and the relationships with the target groups of the evaluated intervention. At the end, I had to take over responsibility for triangulating the surveys, completing the evaluation and presenting and defending its results.

Team leaders are responsible for the results and must be responsible for selection of their team.

Partnership with Local Partners

Does the core team work with local partners (experts, Indigenous people, target groups of development interventions)? Does the core team trust them and vice versa? Do they consider each other's concerns? Do they discuss and correct identified mistakes?

Many commissioners have introduced special budget lines for junior evaluators and local experts. This is good practice; experienced evaluators work as mentors and take responsibility for coordination of an evaluation while young and emerging evaluators perform a significant part of the work. Training by doing and mutual learning are the best ways to test, create and enhance capacities, including the capacities of evaluation leaders.

Case 7: I became a member of IDEAS in 2003, at the International Program for Development Evaluation Training (IPDET) in Ottawa. For four weeks, we had opportunities to work and enjoy together (Play hard, work hard!) and create strong friendships. Thanks to this experience and to all further joint events and working assignments, I can always ask my IPDET friends for personal recommendation of the best people for a concrete evaluation in their country, and vice versa, I am happy if I can work for my friends or recommend a person who fits their needs better than I do.

Personal experience matters much more than curricula vitae.

We have all probably received an e-mail asking for a curriculum vitae two days before a deadline because the terms of reference requested local (or international) experts. Like for independent selection of evaluation team leaders and team members by evaluation clients, it is a lottery, and because of time limitations, there is usually weak ownership of the evaluation design by individual team members. On the other hand, it might be an opportunity to recognize new people and learn from them.

Case 8: I remember the first consultations on country-led evaluation systems in 2005/06. Nowadays, all donors encourage country-led evaluations, but there are still challenges related to capacity development or use of evaluations for transformational change at the national level. Quite often, donors are still in the driver's seat.

Nobody can learn driving from the back seat.

Partnership with Ordering Parties

Is there sincere cooperation between the evaluators and the ordering parties? Are they aware of each other's visions and concerns? Do they share and correct mistakes? Do they respond to recommendations and lessons learned?

Administrative barriers, public procurement rules and hypothetical or real conflicts of interest complicate partnership with ordering parties, especially with public bodies. On the other hand, there is a common goal – to bring reliable evidence and reasoned arguments for improving or expanding results of development interventions (projects, programmes, strategies).

Therefore, close cooperation is necessary and cannot be built merely on client and supplier relations.

Case 9: The Czech Evaluation Society, in collaboration with the Ministry of Regional Development, prepared Guidance on Contracting the Evaluations (CES 2018). Our Code of Ethics (CES 2011) and Standards for Conducting Evaluations (CES 2013) are referenced in most open tenders on evaluations.

We are engaged in the Reference Group on Evaluations of the Ministry of Foreign Affairs and helped introduce the evaluation system for international development cooperation. Several representatives of public bodies are active members of our association.

Mutual cooperation of commissioners and evaluators significantly improves the evaluation culture and the impact of evaluations.

Partnership with Other Evaluators

Does the partnership work between diverse evaluators from the private, public, academic and non-profit sectors? Are they candid enough? Do they speak and listen to each other? Do they support and defend each other? Do they share their experience and learn from each other? Do they cooperate in national or international associations?

There are two contradictory factors – the evaluators are competitors, but they need increasing demand for evaluations.

Case 10: A few years ago, within the Czech Evaluation Society, we conducted a voluntary peer review of evaluations completed in the previous three years. This was a very useful test of the applied standards on conducting evaluations. Later, we became engaged in reference groups or in mentoring for new or ongoing evaluations. These activities are valuable not only for mutual learning, but also for proving integrity and unbiased approaches.

Voluntary participation in peer reviews is valid proof of an evaluator's self-confidence and responsibility.

The quality of evaluations and the use of evaluation results should be a common goal of competing evaluators, across companies and sectors. To

reach this goal, we need more effective cooperation among public bodies, the academic sector and evaluation practitioners; more lecturers or mentors ready to share their expertise and mutual trust.

Opinion 2: Methodological and advisory work, conferences, training events, consultations and mentoring contribute to a better evaluation culture, but active engagement of many more actors is still needed. Many evaluators proclaim their expertise but do not participate in evaluation events or advocacy efforts.

Active engagement in national and international evaluation associations should be considered a key competency of professional evaluators.

Partnerships with Newcomers

How open are evaluators to newcomers – students and interns, young and emerging evaluators from diverse sectors? Do they engage them in evaluations? How honest are the newcomers? How responsible are they for reaching the best results and protecting the whole team (and the target groups)?

Creating and nurturing capacities requires strong, predictable national and international support and enough time for testing and learning. Evaluation is a long-term process, requiring not only expertise and true commitment from all actors, but also sensibility and empathy.

Many young people are activists in diverse sectors (e.g. environmental protection and climate), but they often use 'recycled' arguments from secondary sources (including fake news) or from populist leaders. Engagement in evaluations is a great opportunity to triangulate their sources and methods, discuss the opinions and problems of other stakeholders and identify new ways to contribute to transformation in their countries or communities. At the same time, evaluations need new views, new expertise, innovative approaches and technologies, and especially, engagement of people who care about the future of their societies and are ready to devote their energy to influence the future.

Case 11: In the Czech Republic, we have started the fourth year of an evaluation competition for university students. The winners are invited to international competition and receive vouchers to special evaluation trainings, and some have been working as interns or experts in concrete evaluations.

On the other hand, there is considerable fluctuation of interns in evaluation teams, not only because of a lack of funds, but also of time to create a sense of belonging and solidarity.

Engaging young and emerging evaluators is challenging but brings mutual benefits for both sides.

Specific Aspects of Partnership

Long-Term Partnerships

The COVID-19 pandemic disclosed more clearly than ever before that we cannot succeed alone, in a 'quarantine' of closed communities. We depend on each other. We must be able to nurture contacts and long-term partnerships with other development actors and not just ask for their opinion in one-shot surveys or engage them in one-off contracts when we need them to succeed in a tender. We must 'touch and feel' our findings, not just re-interpret them from statistics. This cannot be ensured using remote sensing or sophisticated teleconferences. Many key respondents do not speak our language, some have poor Internet connections and some belong to marginalized or remote groups. Moreover, body language is much more important than answers in an interview or a questionnaire. For discovering real motivations, causes of problems or overall context, we need reliable team members, facilitators and interpreters in the field. Such people are key assets of stable evaluation teams and cannot be drawn by random sampling in last-minute calls. Teams established on the principles of long-term partnership usually achieve better results.

Case 12: It often takes several failed attempts to find a facilitator or interpreter who understands our mission and can read between the lines. These people are invaluable members of the team.

I remember many cases when body language and eye contact helped reveal key challenges. Online communication cannot replace this experience. I also remember cases when using an unreliable expert or biased interpreter destroyed several months of effort of the whole team.

Recognizing the right people and their true motivations is the most important part of the evaluation profession. Likewise, evaluation is a great opportunity for recognizing the right people.

Working with people is the best incentive and the best reward of the evaluation profession.

Mixed Evaluation Teams

Routine approaches can miss important signals, so we must be open to including new people, unbiased researchers and observers in our evaluation teams. Without feedback and skills from newcomers, 'outsiders' and amateurs, our evaluations would stay old-fashioned, would not reach the right people in the right way and would not reflect real-life and emerging challenges.

Opinion 3: Asking the right questions is an art. Asking the right questions in the right way is difficult. Operational blindness often conceals the fact that respondents do not understand our standard evaluation questions or approaches, which thus cannot bring the hoped-for results.

Business as usual cannot effectively identify the challenges of rapidly changing circumstances.

Associations of Evaluators

Voluntary organizations for professional evaluations can include national, regional, sectoral, international and global associations of evaluators. They are essential for improving and defending the evaluation culture. Their influence must have internal and external dimensions. They must protect their members and create an enabling environment for them, but they must also guarantee their professional integrity. This is not easy.

Everything depends on people. Even a few devoted people can contribute to important changes, and a few unfair people can destroy long-term efforts.

Case 13: An evaluator wanted a new position. When he did not succeed because he violated the election rules, he falsely accused the competitors and the whole organization of racism and personal revenge and made many other allegations.

People who lie or attack other people because of their own hidden interests cannot be unbiased evaluators. All professional associations must have an effective mechanism to defend their ethics.

Cooperation of Associations

Especially at the foundation or transition stage, national and regional evaluation societies need assistance from their international peers. Sharing experience, experts, and advisory and moral support and cooperating on events are a few easy steps that can be decisive for the future of evaluations in a region. Several associations connect evaluation networks from different countries or structures (e.g. Network of Evaluation Societies in Europe, African Evaluation Association, International Organization for Cooperation in Evaluation, IDEAS).

Case 14: IDEAS significantly helped us in starting the European Program for Development Evaluation Training (2007–17) and helped establish the Czech Evaluation Society (2008). Our cooperation continues, and we believe that we support IDEAS as well.

Helping others helps us.

Engaging People

It is necessary to bring evaluations closer to the people. This requires appropriate formulations of key messages and an attractive presentation. Decision makers and target groups (people affected by the evaluated intervention or evaluation results) must understand the findings, conclusions and recommendations. All of us face problems with missing translations, too many unexplained acronyms, too complicated text or too many inconsistencies in reports.

Opinion 4: If evaluators are persuaded that their messages are critically important, they must remember these messages and their justification, without looking to notes, long annexes or lists of abbreviations.

The recommendations 'Keep it simple and short' and 'Make it attractive for your audience' are valid for any influential evaluation.

Prague Declaration on Evaluation for Transformative Change

Based on experience from national and international evaluations, the Czech Evaluation Society prepared a draft declaration, 'Together for Change', to identify commitments that are critical for promoting further use of evaluations for sustainable transformational change. This draft influenced the final declaration, as presented above. All 10 points address engaging people in evaluations and, directly or indirectly, the partnership principles, as I show below in my comments.

Re1. Promote transformational evaluation for the Sustainable Development Goals

If evaluations are to help us learn, understand and support changes, all actors must be engaged, including the most marginalized and the most affected by existential threats. At the same time, all evaluators must know how and why their evaluations have or have not contributed to anticipated changes.

Re2. Work in partnership

Partnership of evaluators is a core principle of influential evaluations. Ethics codes, professional standards, mutual trust and engagement of diverse actors with the aim of mutual learning are explicitly mentioned here, but support from and engagement of commissioners and other actors are also necessary.

Re3. Explore power relations and promote inclusiveness

Inclusive evaluation approaches call for engaging local stakeholders and Indigenous people and incorporating local ways of knowing. This requires partnership with these actors because they must be engaged in all stages of evaluations that concern their lives, and they must share responsibility for the results. Evaluations have a strong empowering role in developing local capacities and can help change power relations.

Re4. Respect for rights and responsibilities

All partners must apply codes of ethics and professional standards when conducting evaluations. A proper monitoring mechanism must also be in place.

Re5. Support for professionalization and capacity development

In addition to the professional training events and supporting role of the International Evaluation Academy, the potential for mutual learning for all evaluation stakeholders and cross-sectoral sharing of expertise and experience must be considered. Professionalization of evaluations requires professional feedback and using lessons learned. Voluntary peer learning among partners is a valuable tool.

Re6. Focus on sustainability

Good evaluators must assess all relevant contextual factors and be aware of all pillars and assumptions of sustainable development, including placing people and planet at the centre and leaving no one behind. Cross-sectoral partnerships are necessary in this regard.

Re7. Focus on fragility, conflict and violence

The COVID-19 pandemic has revealed the importance of new aspects of fragility that affect all countries – the importance of supporting community resilience, of strength and solidarity and of engaging local partners.

Re8. Support for transformational indigenous evaluation

The role and expertise of Indigenous people should be reflected in the partnership schemes of evaluations, based on brainstorming and co-creation (see The Partnering Initiative and UN DESA 2020). These partnerships should go beyond the first four levels of 'informing, consulting, involving and collaborating' to the fifth level of 'empowering' (see UN DESA and UNITAR 2020).

Re9. Shared responsibility for results

Mutual cooperation of evaluators and evaluation users is crucial for the use of results and thus determines the usefulness of each evaluation.

Re10. The challenge ahead

Emerging development challenges (including the impacts of the COVID-19 pandemic) call for transformational changes more than ever before. All the unprecedented challenges need partnerships with engagement of all development actors. The evaluators are often in the frontline of these efforts... or should be.

Conclusions

I have a very special personal experience. From 1968 to 1989, I was living in a totalitarian regime under the military occupation of the Soviet Army. I could not do what I wanted, I could not read and listen to what I wanted, I could not study what I wanted, I could not select my profession and I could not travel to non-communist countries. On the other hand, I had the unique opportunity to contribute to the most important transformational change in our region – from communist totalitarianism to open democracy. In my case, it started with petitions and protests in the streets, continued with establishing a private company and several non-profit organizations and then with contributions to several national strategies and international guidelines and has culminated with my engagement in international development cooperation, and evaluations in particular.

Case 15: Experience is untransferable.

When my wife and I married in 1985, we asked the State Security (secret police) for a permit to spend our honeymoon week in Yugoslavia. (One had to have a reason even for travel to this Eastern-bloc country, and a honeymoon seemed to be a legitimate reason.) I received the permit, but my wife did not. We stayed home.

The 'Velvet Revolution' in 1989 and the peaceful divorce of the former Czechoslovakia in 1993 cannot be easily replicated. Moreover, some people still hate these transformations. The Soviet Union and Yugoslavia do not exist anymore, yet tensions remain. We have historical evidence of the inhumanity and absurdity of totalitarian regimes at the cost of millions of destroyed lives, but people are still suffering in many regions of the world. All people matter...

Transformation is not an easy process, and it cannot be imposed from outside. Neither can happiness. Evaluations can discover the genuine drivers for change and show the way.

I was living in a country with a totalitarian regime that became an aid recipient and a country in transition and is now a member of the European Union and the Development Assistance Committee of the Organisation for Economic Co-operation and Development. This does not mean that our democracy works perfectly and that our transformation is complete. It will

take at least one more generation to transform the totalitarian mindset of many people. Global turbulence may slow this transformation even further.

This experience is invaluable. I can understand some issues that people who did not live through a similar situation cannot easily understand. I also believe that I have a stronger motivation to contribute to systemic changes than people without transition experience.

Nevertheless, my experience is still limited because I am only human. I can provide neither definite guidance on how true partnerships should work to bring the best results nor the answers to many of the above questions, but asking the questions is a basic evaluation tool. Any opinion, experience or personal concern can be a good start for discussion.

The best solutions usually come from brainstorming and teamwork. Effective solutions then need transformational (systemic) change and the personal commitments and joint efforts of many devoted people from diverse sectors. I can confirm that interdisciplinary and multisectoral partnerships really work, although we usually have diverse and opposing roles. I have worked for or with the state authorities in several countries, the academic sector, non-profit organizations, the private sector, international agencies, financial institutions and Agent Orange victims. All that experience has confirmed the indisputable laws of sustainable development: 'Everything is linked to everything else' and 'Everything depends on people'. If we speak and listen to each other, are fair and empathetic and strive to make a better life for our families, communities, countries and the planet, how much easier everything will be.

I wanted to share several messages – burning questions – from our discussion of the main theme of the 2019 IDEAS Global Assembly: 'Evaluation for Transformative Change: Bringing Experiences of the Global South to the Global North':

- Do we return again to the paradigm of North and South, donors and recipients? Where is the place of the Czech Republic, being a Western country until the Second World War, becoming an Eastern country within the Soviet bloc afterward and then an aid recipient and thus part of the Global South after 1989 and now being a part of the Global North? We did not move; we are still in the centre of Europe.
- How can the Czech Republic and other recent transition countries contribute to development? Although we do not have as much money as more experienced donors, we are not eligible for donor funds anymore, but we have strategic experience of transition, we

succeeded and failed in many transformational efforts (and partially learned from the failures), we have empathy for our development partners and we can improvise quite well. Is it enough?

- What can we do to strengthen partnerships between diverse development actors? Can we draft a declaration that could name the key issues and propose a way forward? Can we contribute to implementation of the agreed-upon commitments? I believe so. The first step – approval of the Prague Declaration – was successfully completed. The second step has started as well – we are providing a country office for administration of IDEAS, and we truly believe that IDEAS can become even more attractive to the global evaluation constituency and more influential. Together, we can make it.

In this chapter, I discussed principles of effective partnerships in evaluations for transformational change. I also explained why I consider the 10 points of the Prague Declaration important and how closely they relate to necessary joint efforts and partnerships of all development actors. To document my opinion and recommendations, I used real-life stories from my career. I am ready to receive any feedback from my peers.

I am proud to be a member of the IDEAS family.

Testimonials

Silvia Salinas Mulder

I am a Bolivian anthropologist, evaluator, innovator, feminist and human rights activist. Bolivia is a country where more than 60 per cent of the population is Indigenous and approximately one-third lives in poverty and where colonial relations persist and shape daily life. It is also a country with one of the highest rates of femicide in Latin America. In addition, Bolivia has been rated as the most distrustful country in the region.

My multiple interlinked identities, the reality of my country and my self-reflection and learning processes have influenced my career as evaluation practitioner, activist and leader. To be honest, I do not clearly identify how or when I became an evaluator, because like most Bolivian evaluators, I am self-taught, but I am absolutely sure that it had to do with my ambition to contribute to making the world a better place for everyone.

I find the Prague Declaration very relevant, and it resonates for me as an invitation to individual self-reflection and change. I strongly believe that the world – and consequently evaluation – will not change unless we all start taking charge of our ways of thinking and doing, our attitudes and beliefs. Can we challenge ourselves and transform the Prague Declaration into an individual self-assessment?

Power, rights and inclusion are at the centre of the Declaration. Although some statements are dedicated specifically to them, I think that their implications are especially relevant for all statements and are at the heart of the role of evaluation in 'promoting learning, systemic and transformational change'. COVID-19 has reaffirmed the urgent need for profound systemic change in our human paradigms. It has also confirmed the potential role of evaluation in guiding those transformations; the need for change in how we think, do and use evaluation is also evident and is a necessary condition for its potential to be unveiled. This implies questioning the global evaluation architecture and the assumptions that govern the understanding, relations, decisions and budgets, which tend to be Northern, Western, adult and male-biased.

In recent years, a group of female evaluators from the Global South dedicated ourselves to influencing international dialogues and opening a discussion about the nature and underlying power relations of the international evaluation architecture and agenda. We advocate for a South–North horizontal dialogue that enables the evaluation paradigms to be reinvented

while recognizing the contributions and rights of the Global South. There are many yet unveiled and taboo issues in the evaluation arena that we need to address in a frank, open dialogue; as the Prague Declaration states, 'we continue to discuss and to deepen our understanding of the changes required for evaluation to contribute to tackling the crucial problems of our time'.

Partnerships and collaboration, addressed in statement 2, are vital to creating a viable future but are only possible if we recognize and respect diversity and are capable of focusing on complementarity beyond affinity. We must also understand inequality and recognize our co-responsibility in the reproduction of multiple biases, discriminatory behaviours and hierarchical power relations.

In recent weeks, we were challenged to find 30 Latin American, English-speaking Indigenous evaluators to be sponsored to attend the upcoming Conference of the Canadian Evaluation Society. Despite our efforts, we were unsuccessful. Indigenous people are not expected to be evaluators but evaluated 'beneficiaries', and the few that have managed to overcome the discriminatory structures do not speak English. This encourages us to apply the 'no one left behind' Sustainable Development Goal mandate in the evaluation field, addressing exclusion factors such as education and language and transforming the systems and relations. On a personal level, although I commit to statement 8, 'to value and support the strengthening of and learning from indigenous evaluation by and for Indigenous peoples', I must recognize that, as a white, urban evaluator, I have reproduced colonial, paternalist relations with Indigenous rural female 'beneficiaries'. Even my awareness and good will are not enough to change history, perceptions and centuries of colonial mindsets and relations. From another perspective that contributes to understanding the complexity and multifaceted nature of power relations, I also recognize that being female has implied facing sexist attitudes and even disrespect of male rural Indigenous authorities.

Power and ethics are interrelated. In 2000, I published with other colleagues an article titled 'Unethical ethics?' addressing experiences and reflections in intercultural research practices. I think that 'unethical ethics' is common in evaluation; I have witnessed lack of respect, sexism, colonialism and other types of unethical, more or less explicit behaviour in evaluations, with no implications. Despite advancements in establishment of ethical codes and procedures, ethical compliance often remains as a formal aspect that does not penetrate the 'evaluation DNA'; it is not integrated into the evaluation activities, relations and organizations, and we seldom reflect on it.

Self-reflection, from my perspective, is a critical but not frequently considered evaluation competency, probably because it also relates to the

idea of being humble...and this contradicts conventional ideas about our role and power position. In 2015, I was part of an exercise to compare three evaluation competency profiles. The main finding was that the ideas about the desired evaluation competencies, the underlying assumptions and the image of a good evaluator differed dramatically. Although human-centred competencies have increasingly been considered in evaluator competency profiles in recent years, professionalization and certification programmes still do not pay sufficient attention to competencies needed to address power, ethics, diversity, gender and inclusion in evaluation practice.

My final reflection relates to the voluntary organizations for professional evaluation and in general to the different organizations, partnerships, coalitions and initiatives that shape the rich, although complex and competitive, global evaluation ecosystem. My recent experience as chair of the regional Monitoring, Evaluation and Systematization Network of Latin America and the (Spanish-speaking) Caribbean, president of the International Organization for Professional Evaluation since 2020 and co-chair of EvalPartners starting in January 2021 has reinforced my idea that organizations must practice what they preach; we must all walk the talk and make our own Prague Declaration self-assessment!

Rashmi Agrawal

The Prague Declaration is a succinct expression of intent on the part of all partners in development to promote and use evaluation as a tool to bring about the transformational changes needed to achieve the Sustainable Development Goals. Transformational changes need, apart from material resources, behavioural changes that bring about universal respect for the environment, a healthy life and learning. It is important, therefore, that evaluations look closely at assessing desirable changes in mindsets and that evaluators look closely at effective means of such assessment. Personally, I have always been fascinated by qualitative methods that directly involve participation of stakeholders at all stages of evaluation have always fascinated me. I have been, for instance, arguing for story-telling by and analysis of the narrations of participants as an approach that holds considerable promise. I had, at an earlier IDEAS Global Assembly in Guanajuato, introduced this approach in a pre-conference workshop.

Evaluation of transformational change requires new and systemic approaches. Dissemination of this newly acquired knowledge using wide-spread initiatives to develop national evaluation capacities is of the utmost

importance. Through the activities of the Evaluation Community of India (ECOI), an association of professional evaluators of which I am a founding member and continue to be part of a core group managing its affairs, we have been pursuing this goal over the past five years. Our approach has been to encourage emerging evaluators to innovate and share their products with a wide range of stakeholders. The Innovation Bazaars, organized as a part of our EvalFests (event-facilitating meeting of stakeholders in evaluation) in 2018 and again in 2020, have attracted many ideas and much talent from young and emerging evaluators. In all these events, we have provided ample space for youth, and the opportunity was used with excellent results. The launch of EvalYouth India Chapter as part of ECOI has opened opportunities for participation of young and emerging evaluators (YEEs) in helping generate evidence for decision-making. A collaborative initiative of the APEA along with other associations of evaluators was the Asia Pacific Virtual Winter School 2021 for YEEs.

ECOI did not look only at new entrants to the evaluation profession. My interactions with a Delhi college (India) indicated immense interest on the part of the students to learn the basics of monitoring and evaluation, beyond their regular curriculum. In an extension of our efforts to cater to this emerging need, in collaboration with the faculty and administration of the college, we organized training of its students in the faculty of management studies in the principles of monitoring and evaluation. We intend to continue and expand our efforts in this area.

The onslaught of the COVID-19 pandemic has not deterred the evaluation community from evaluations or evaluation capacity development. As the pandemic posed immense challenges to the application of usual methods of generating evaluative evidence, resilience on the part of the evaluation profession opened doors for newer approaches to data collection and transmission of knowledge. An exercise that I have personally undertaken with a few friends assessed the psychological effects of the prolonged lockdowns and the changed life and work styles on people using web-based surveys. Increased use of technology in evaluations has widened the range of information available for decision-making.

The Prague Declaration emphasizes working in partnerships between evaluation stakeholders. My engagement in partnerships did not end with learning interactions with young evaluators and students. We have forged collaborative partnerships with other voluntary organizations of professional evaluators. For example, ECOI has entered into memoranda of understanding with the Asia Pacific Evaluation Association, Sri Lankan Evaluation Association, Indonesian Development Evaluation Community

and Afghan Evaluation Association. These partnerships have yielded rich dividends, particularly in evaluation capacity development. A series of joint webinars on wide-ranging topics of current relevance has helped evaluators in these countries learn a lot. Some of the topics covered included those relevant for transformational change, such as evaluation of climate change, Blue Marble Evaluation and gender-focused evaluations.

Parliamentarians hold the key for demand for evaluation and use of its results. ECOI has therefore engaged with legislators in its deliberations on various topics connected with evaluation in EvalFest 2020. We intend to carry forward this mutually beneficial dialogue to create an evaluation-friendly eco-system in the country. A similar initiative is the partnership with state agencies. A statement of intent has recently been signed with the Development Monitoring and Evaluation Office, the nodal agency for monitoring and evaluation in the national government, with ECOI and a few other organizations to strengthen the monitoring and evaluation system in the country.

An immediate follow-up to this partnership has been the participation of ECOI in the National Conference on Monitoring, Evaluation and Learning that the National Institution for Transforming India organized. A panel presentation that ECOI sponsored at this conference that I moderated addressed professionalization of evaluation. ECOI has also been partnering with the Asia Pacific Evaluation Association and several other voluntary organizations of professional evaluators in the Inter-Regional Initiative for Professionalization of Evaluation, components of which include developing a definition of professionalization, competency frameworks and ethical standards. The outcomes of this initiative will feed into the work on professionalization proposed in the Prague Declaration that the International Evaluation Academy will take up. I was a member of the panel that discussed the need for such an academy at the Prague Global Assembly of IDEAS.

We hope that this momentum in strengthening the demand and supply sides of evaluation will enhance the quality of evaluations and their relevance for evaluating transformational changes and lead to a greater degree of evidence-based decision-making that would help in achieving the Sustainable Development Goals.

Josephine Watera

The IDEAS Global Assembly, held in Prague 2019, was a successful, highly memorable event. I was excited to attend the conference. The theme, the discussions, the keynote speakers and the choice of location offered nothing but the best. The highlight of the assembly was the adoption of the Prague Declaration on Evaluation for Transformational change. This milestone proved that, when we come together, we grow and develop together. The declaration calls for commitments from all well-meaning individuals, organizations and societies to advance the discourse of evaluation for sustainable change.

As an evaluation practitioner in the parliament of Uganda, I observed that parliamentarians more often than not address crises that urgently require evaluation information and action: epidemics; pandemics and socioeconomic, environmental and political crises. The Prague Declaration reminds me that I must be intentional in promoting transformational evaluation for realization of the 2030 Agenda for Sustainable Development. Partnership is specifically of great importance to me today, considering the great contribution of professional bodies, civil society organizations and academia in generating evidence to inform transformational change in Uganda. These bodies largely advance expert knowledge and objective research, which is critical for this declaration. Conversely, committees of parliament offer an open platform for supplying evidence, but evaluators do not always use this opportunity; even when evaluators appear before committees of parliament, the mode of communicating evaluations is still highly technical and not very usable for quick decision-making. The Prague Declaration's call for support for professionalization and capacity development seeks to address this gap; hence all efforts must be put in place to make the International Academy a reality and translate the same efforts at regional and national levels.

In Uganda, there is a strong focus on local content, which is in line with the third article of the Prague Declaration: designing evaluations that include Indigenous and local ways of knowing with conventional and transformational methodologies. This shift in mindset has significantly changed how my role is perceived and attention given to evaluations containing Indigenous 'know how' methodologies. With such efforts of knowledge transfer and promotion of Indigenous knowledge, the journey towards sustainability, with specific focus on contextual and system perspectives, seems ensured. The Prague Declaration commits to valuing and supporting

the strengthening of and learning from Indigenous evaluations by and for Indigenous people.

As the Prague Declaration states, I concur that the real result of an evaluation is not the evaluation itself but the use that is made of the evaluation. I have played the role of knowledge broker to bridge the information gap for evaluations to inform the business of parliament. There are targeted efforts to ensure that evaluations are available in a timely, usable manner. Specific efforts in place are: a framework for civil society participation in parliament business, open parliaments with access to live streaming of parliamentary sessions, access to the order paper (daily agenda of parliament) and annual parliamentary calendars. This has helped evaluators know what is happening in parliament, what evaluation is needed and how to use it in decision-making.

The Prague Declaration recognizes the challenge of awareness of and understanding the discourse on evaluation for transformational change in some working systems. This is not different in Uganda, where the field of evaluation is still growing, and even more critical in parliament. The Declaration is a continuous reminder of my role and the role of other practitioners in increasing awareness at every opportunity.

The Prague Declaration has therefore proved not just timely, but also a focused and very relevant instrument of global transformational change. If all of us can draw our energies and commitments towards this Declaration, we shall live to see the change, the transformational change, for current and future generations drawn from evaluations. What a world this would be.

Statements

Ada Ocampo, President of IDEAS

The World Health Organization declared the COVID-19 outbreak a Public Health Emergency of International Concern on 30 January 2020 and a pandemic on 11 March 2020. The pandemic has had negative consequences in all dimensions of societies. It has changed our lives, as well as how we do business. Between March and July 2020, organizations and experts wrote about and embarked on continuous discussions about the need to approach evaluation differently during the pandemic. Proposals for new approaches and ways of working rapidly emerged. Vibrant discussions led to more questions than answers. In the midst of a crisis during which fear and uncertainty permeated our lives, there was an opportunity to review evaluation and to discuss and agree on ways to ensure that evaluation as a function and a profession will remain relevant. This period was very exciting. I was elected President of IDEAS during this period. Although I was eager to start my new position, the pandemic forced me to remain in New York for much longer than expected. I was not able to join the IDEAS Board formally until 2021.

The pandemic posed similar challenges to the one I faced to evaluators throughout the world, especially evaluators working internationally and evaluations taking place in several countries and regions. It was no longer possible to organize field visits for all team members or to have physical meetings of the team to discuss the evaluation and, perhaps more importantly, to meet with policymakers, stakeholders and local communities or to organize focus group meetings. This sudden and unexpected challenge was met with high degrees of improvisation and discovery of new ways of working together, and we should applaud the ingenuity and resourcefulness of evaluators in ensuring the quality, relevance and usefulness of their work in these circumstances.

The pandemic thus put the response of the international evaluation community to the Prague Declaration in a new and unexpected light. IDEAS has been the birthplace of the Prague Declaration, but it was facing new challenges when carrying this forward. Follow-up actions to the Prague Declaration were also delayed or changed in nature. A Wilton Park meeting on transformational change was foreseen in July 2020 but had to be postponed. This meeting in turn was supposed to provide input for the negotiations for climate action at the 26th United Nations Climate Change Conference of the Parties of the United Nations Framework Convention for Climate Change, expected to take place in November 2020 but postponed until November 2021.

By necessity, the follow-up work based on the Prague Declaration has been virtual, in writing, as perhaps best expressed in this book. For IDEAS, a key element in transforming evaluation to support and strengthen transformational change is professional capacity development. IDEAS has worked on a framework for professional competences for evaluators, managers and commissioners and has adopted a code of ethics, both finalized in 2012–2014, before transformational change became a rallying call in Agenda 2030 (United Nations 2015). Together with other partners, IDEAS must update these documents and make them relevant for our times. I see this as an important agenda item for IDEAS to take up.

Furthermore, as the Prague Declaration states in various places, we must act together, which means enhancing our approach to partnerships, as Daniel Svoboda so ably voiced in his statement in this chapter. One of the organizations coming directly out of the Prague Declaration is the International Evaluation Academy, as proposed by IDEAS. Keeping aligned with the Prague Declaration, we are embarking on strategic joint ventures with EvalPartners, the United Nations Children’s Fund, the International Fund for Agricultural Development, the Independent Evaluation Group of the World Bank and others. The International Organization for Professional Evaluation, the umbrella organization of voluntary organizations for professional evaluations, and IDEAS, the only global professional association for international evaluators, are increasingly cooperating on evaluation challenges. Last but not perhaps most important, IDEAS has members in many countries and will continue to aim to be relevant to them and support them in their struggles to increase their capacities in national monitoring and evaluation systems and on the ground. We look forward to continuing to work together to meet the challenges of transformational evaluation.

Juha I. Uitto, Director, Independent Evaluation Office of the Global Environment Facility

Much has happened in the relatively short time since the adoption of the Prague Declaration on Evaluation for Transformational Change in October 2019. The COVID-19 pandemic broke out just a few months later and turned the world upside down, devastating lives and wreaking havoc on economies. Meanwhile, climate change has continued unmitigated, its impacts becoming increasingly clear in our everyday lives as hurricanes, wildfires and weather anomalies add to societal stress around the world. We continue to lose biological diversity and valuable ecosystems at unprecedented rates. Poor people and minorities are especially vulnerable to the

impacts of the pandemic and environmental changes. Societal polarization is greater than in decades. All of these trends underscore the importance of what the Prague Declaration advocates for evaluation.

The pandemic, at its root, is an environmental crisis. The virus causing COVID-19 is zoonotic, meaning it has crossed to humans from non-human animals. Such spill-overs are increasingly common – and dangerous – because of how humanity infringes on the natural environment. Research clearly shows how deforestation and habitat destruction favour disease-transmitting species (e.g. rats and bats) and bring them into ever-closer contact with people. Globalization and rapid movement of people facilitate the spread of pathogens. Wildlife trade is another factor that destroys the environment and poses a hazard to human health. An evaluation of the Global Environment Facility's Illegal Wildlife Trade programme by my office highlights the need to take a comprehensive approach to such problems. Promoting local livelihoods in source countries to discourage poaching is important, but it is also necessary to address demand in destination countries in Asia, Europe and North America and to address enforcement and corruption throughout the transit chains. Evaluating such complex programmes requires an inclusive perspective, varied approaches, knowledge and partnerships.

The sixth principle of the Prague Declaration commits evaluators to focus on social, environmental and economic sustainability and transformation, given the close interlinkages between the three dimensions that also underlie the 2030 Agenda for Sustainable Development. The pandemic, as well as climate change, have highlighted that we humans are part of the broader ecosystem and that environmental health and human health are intertwined.

Another important principle in the Declaration is its seventh point: focus on fragility, conflict and violence. This is an area about which we at the Global Environment Facility (GEF) must learn more. As our recent evaluation of GEF support in fragile and conflict-affected situations unequivocally demonstrates, these characteristics influence programme and project performance and sustainability through various pathways. These are also situations in which people are the most vulnerable, not only to political, security and economic risks, but also to the impacts of environmental change.

Conducting a comprehensive evaluation of the GEF to feed into the quadrennial intergovernmental negotiations to replenish the fund has had its challenges during the pandemic. The comprehensive evaluation consists of 34 separate component evaluations and studies, ranging from the

impacts of the GEF's various programmes to organizational effectiveness. As always, collecting evidence from the field is critical. Fieldwork had to be put on hold when the pandemic hit, but we needed the perspectives of the governments, civil society and the people intended to benefit from the GEF interventions. We responded by engaging our network of consultants living in partner countries who could still conduct field visits and interviews safely. At the same time, we employed tools such as remote sensing and geospatial analysis to detect changes in the natural environment, land use and other variables that could be tracked remotely. Both approaches were successful and demonstrated that serious evaluations can be conducted this way and with a smaller environmental footprint for the evaluation itself.

At the Independent Evaluation Office, our evaluation practice reflects the values and principles embedded in the Prague Declaration. Our goal is to bring evaluative evidence for learning and for promoting systemic and transformative change for the benefit of the global environment and we people who depend entirely on the health of the planet. Evaluating at the nexus of natural and human systems is an area where I think the evaluation community still has much to learn. Taking the principles of the Prague Declaration to heart is a good start.

Rob D. van den Berg, former President of IDEAS (2014–2020)

A long time ago, when I started as Director of the Policy and Programme Evaluation Department in the Dutch Ministry of Foreign Affairs, I had the great experience of seeing an evaluation in my department come up with a highly relevant insight into how global change in some cases is initiated and takes shape. This was long before we called fundamental changes 'transformational'. This evaluation was focused on new institutional perspectives on sustainable management of water resources. During the 1980s, the view gradually emerged that an integrated approach was necessary that would not just deliver drinking water to households, but also include management of groundwater resources, drainage, irrigation support and sanitation. This integrated perspective was agreed upon internationally at international conferences in Dublin (1992) and the Earth Summit in Rio de Janeiro (1992).

In 2000, when the evaluation was published, it concluded that the new, integrated perspective, which included social, economic and environmental issues, took on average five years, from 1992 to 1997, to become visible in adopted policies in countries; this average was the same for countries in the Global North as for those in the Global South. Early adopters of

the new concept had a similar time gap of five years between approving a new policy and showing changes on the ground. Upon reflection, this was a transformational change in management of the water sector that took years to develop into an international agreement, reached in 1992 in Rio de Janeiro, and then took on average of five years to seep through into national legislation and budgeting, leading to more years before the new policies were visible on the ground.

The global crises of our times, as in the title of this book, are addressed in only a scattered and fragmented body of international agreements, and where agreements have been reached, they tend to be aspirational rather than concrete, underfunded rather than fully budgeted and without fully coordinated international action. An example is the global response to COVID-19; countries were often thrown back on their own resources. Borders were closed. Competition for medical equipment and development of vaccines was rife. President Trump found the pandemic a good reason to stop U.S. funding of the World Health Organization. President Biden has turned this around, but the rise of populism in the world does not bode well for international cooperation and action.

A relatively small, but ever growing, group of people is fully discussing and endorsing transformational change for the Sustainable Development Goals (SDGs) and for climate action. IDEAS is a good example of how long it takes for new ideas to capture the imagination and lead to new paradigms and, above all, action. When I became president of IDEAS in 2014, the Sustainable Development Goals were being drafted. Reading these drafts, many of us thought the United Nations would never agree to these goals. They were too aspirational, too integrated and perhaps most importantly, too transformational. For a short while, we lived in a dream world, when in September 2015, the United Nations unanimously adopted Agenda 2030, which starts with the rallying call to 'transform our world' and includes the SDGs in full – not in a watered down version, but with all the transformational and system perspectives included.

IDEAS prepared a Global Assembly in Bangkok in November 2015, and while this preparation was taking place, we did not have any clear perspective on how Agenda 2030 would take shape. We decided to focus on sustainability – the underlying concept of the SDGs and the underlying reason for the global crises of our times. Looking back, one may wonder why we did not focus on transformational change and how evaluation could support it in November 2015. In other words: why wait until October 2019, at our Prague Conference? The reason is the slow maturation of ideas and concepts; they take time, even when there is hardly any time left. In

2015, the international evaluation community was not yet fully focused on systems and thought of transformational change as something they were not involved in, and many had hardly heard of complexity science, systems thinking or non-linear, chaotic, risky developments in the real world.

This book is testimony to what we have learned since 2015. The IDEAS publications of 2017, 2019 and now this volume show a transformational perspective on the challenges of our times and the role of evaluation. Although the first publication, *Evaluation for Agenda 2030* (Van den Berg et al. 2017), moved in the direction of highlighting regional and national Southern perspectives, it covered new ground, such as impact investing and sustainability of impact, in only a few chapters. The second publication, *Evaluation for Transformational Change* (Van den Berg et al. 2019), explored new ways of approaching evaluation, from Osvaldo Feinstein's proposal for dynamic evaluation to transformational evaluation in the Global South, value-based evaluations, lessons from the environmental funds and systems thinking in evaluation. This publication increases that diversity and aims to inspire evaluative action for transformational change, because the global crises of our times demand it.

It has been a personal honour to be involved in this voyage from initial recognition of aspirations, without a full understanding of what this means, to a wide array of chapters that show the full range of what is needed, coming from a broad spectrum of writers, from experienced to young and emerging, from all regions of the world, including Indigenous perspectives, leading to a smorgasbord of inspiring and aspiring approaches, ethics, methods and tools, as well as institutional thinking of how this could become a reality. One may hope that, while the world is slowly breaking free of the clutches of the COVID-19 pandemic, this book may function as a source for rethinking and transforming evaluation to better serve the world.

References

- Aronsson, Inga-Lill, and Hur Hassnain. 2019. Value-based evaluations for transformative change. In: *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*, (eds.) Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder. Exeter, UK: IDEAS.
- CES (Czech Evaluation Society). 2011. Code of ethics. https://czecheval.cz/Dokumenty%20-%20archiv/ces_eticky_kodex_en.pdf
- CES (Czech Evaluation Society). 2013. Standards for conducting evaluations. <https://czecheval.cz/en/Documents>

- CES (Czech Evaluation Society). 2018. Guidance on contracting the evaluations. <https://czecheval.cz/cs/Aktivity/Methodika-pro-zadavani-evaluaci>
- GPEDC (Global Partnership for Effective Development Co-operation). 2012. Global Partnership for Effective Development Cooperation. <https://tinyurl.com/2je9a9td>
- IDEAS (International Development Evaluation Association). 2012. IDEAS competencies for development evaluation evaluators, managers, and commissioners. <https://ideas-global.org/policies/>
- IDEAS. 2014. Code of ethics. <https://ideas-global.org/policies/>
- IDEAS. 2019. Prague Declaration on Evaluation for Transformational Change. <https://tinyurl.com/bcm294k7>
- OECD DAC (Organisation for Economic Co-operation and Development's Development Assistance Committee). n.d. The Paris Declaration on Aid Effectiveness (2005) and the Accra Agenda for Action (2008). <https://tinyurl.com/39khfw4b>
- The Partnering Initiative and UN DESA (United Nations Department of Economic and Social Affairs). 2020. The SDG Partnership Guidebook: A practical guide to building high impact multi-stakeholder partnerships for the Sustainable Development Goals. Darian Stibbe and Dave Prescott, The Partnering Initiative and UN DESA.
- UN DESA and UNITAR (United Nations Institute for Training and Research). 2020. Stakeholder engagement & the 2030 agenda. A practical guide. UN DESA and UNITAR, Geneva, Switzerland.
- United Nations. 2005. General Assembly resolution 60/214. Towards global partnerships: Enhanced cooperation between the United Nations and all relevant partners, in particular the private sector. <https://tinyurl.com/3ntxjhxz>
- United Nations. 2015. Sustainable Development Goals, Goal 17 – Strengthen the means of implementation and revitalize the global partnership for sustainable development. <https://tinyurl.com/296u6e76>
- United Nations University Institute for the Advanced Study of Sustainability. 2018. Partnering for sustainable development – Guidelines for multi-stakeholder partnerships to implement the 2030 Agenda in Asia and the Pacific. Tokyo: United Nations University Institute for the Advanced Study of Sustainability.
- Van den Berg, Rob D., Indran Naidoo and Susan D. Tamondong (eds.) 2017. *Evaluation for Agenda 2030. Providing evidence on progress and sustainability*. Exeter, UK: IDEAS.
- Van den Berg, Rob D., Cristina Magro and Silvia Salinas Mulder (eds.) 2019. *Evaluation for transformational change: Opportunities and challenges for the Sustainable Development Goals*. Exeter, UK: IDEAS.
- World Bank. 2004. *Evaluation and development – The partnership dimension*. World Bank series on evaluation and development. Washington, DC: World Bank.

Contributors

Mehjabeen Abidi-Habib is a Senior Consultant with Emerald Network Ltd. Originally from Pakistan, his first evaluation was a few decades ago when a female British forester and he evaluated a male-dominated forestry institute and suggested ways to create a place for the training of the first female foresters. Now he undertakes international evaluation work, and the local perspective is his recurrent theme to gain insight for larger governance scales. Another continuing theme of interest is understanding context and history as ways of placing equitable adaptation efforts. He holds a PhD in socio-ecological resilience based on how local communities face shock and surprise in their ecosystems. mehjabeen.emeraldnetwork@gmail.com

Marie-Hélène Adrien is a Senior Associate Consultant with Universal Management Group in Montreal, Canada. She has 30 years of evaluation experience as a practitioner, professor and researcher. Her particular expertise lies in evaluating organizations and partnerships funded by private philanthropies and multilateral and bilateral agencies in the Global North and Global South. She was the first woman elected President of IDEAS (2005–2008) and serves on the IDEAS board representing North America and supporting EvalPartners. Marie-Hélène mentors young and emerging evaluators in their professional journey and advocates for the use of evidence by policymakers. She specializes in education and holds a PhD from McGill University, Canada. mhadrien@gmail.com

Rashmi Agrawal is a consultant in monitoring and evaluation; previously, she was a Director with the government of India. She has three decades of experience as a trainer and researcher in the field of monitoring and evaluation and other social issues. She has authored a number of books on social issues

and presented several papers in national and international conferences. She initiated and designed a diploma course in monitoring and evaluation as course director that was launched for international participants and had the approval of the government of India. She has served on the IDEAS Board and is a founding member of the South Asia Community of Evaluators (CoE-SA). Additionally, she initiated and is a core group member of the Evaluation Community of India (ECOI), a member of EvalGender+ and a board member of the Asia Pacific Evaluation Association (APEA). She has a PhD in psychology from Lucknow University, India, with a specialization in rehabilitation and counselling from San Jose State University, United States. She is also an International Program for Development Evaluation Training (IPDET) graduate. rashmi_agrawal56@rediffmail.com

Inga-Lill Aronsson holds a PhD in cultural anthropology and is currently Senior Lecturer in Museum and Heritage Studies at Uppsala University. She has extensive experience with multidisciplinary research, teaching, supervision and curriculum development. She merges two fields of inquiry – heritage and development – in investigating memories, heritage and reconciliation in post-conflict and disaster areas, and longitudinal and evaluation research in resettlement. Aronsson has conducted classic ethnographic field work, particularly in Mexico, and has been a Visiting Professor at several universities around the world. She is engaged in debates on academic freedom and ethics and has published a variety of scholarly research pieces, alone or in co-authorship. Aronsson is the former Director of the Network of Humanitarian Action. At IDEAS, she is the elected two-term board representative for Europe and the coordinator of the Thematic Interest Group on Resettlement. inga-lill.aronsson@abm.uu.se

Nima Bahramalian is an Associate Industrial Development Expert of the United Nations Industrial Development Organization (UNIDO). Before joining UNIDO, he worked as a community economic development planner at the Ministry of Amerindian Affairs of Guyana to support sustainable development of rural indigenous communities. Previously, he worked as a mechanical engineer in the private sector in Iran. Over the past 10 years, his career as a development practitioner has focused on managing technical assistance projects that promote sustainable development of productive sectors in Afghanistan, Albania, Belarus, Guyana, Indonesia, Iran, Kyrgyzstan, Republic of North Macedonia, Pakistan, the Philippines, South Africa, Uganda and Vietnam. Nima holds a BEng from the Islamic Azad University of Science and Research Branch, Iran; an MA in international business from the University

of Greenwich; and an MSc in urban economic development from University College London, United Kingdom. n.bahramalian@unido.org

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 (AEA) since 2006 and was chair of the Mixed Methods Topical Interest Group from 2008 to 2018. She is a former 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 (S2SE), which links the CEI to other voluntary organizations of professional evaluators (VOPEs) in the Global South. lbaptiste.mepro@gmail.com

Kenza Bennani is an Evaluation Consultant with Universal Management Group, specializing in gender equality and women's empowerment. She previously worked as an Evaluation Manager with the French Ministry of Foreign Affairs, Directorate General for International Development. She has 10 years of work experience combining evaluation, management consulting and communications roles. She is engaged in several voluntary organizations of professional evaluators (VOPEs), mainly as a board member of IDEAS (representing the Middle East and North Africa region), as a co-leader of the Francophone Network of young and emerging evaluators (RF-Ee) and as the co-founder and co-convener of the French Evaluation Society thematic working group for young and emerging evaluators (JEEunes). With a multifaceted academic background, she holds graduate degrees in political science, sustainable development geography and business administration. kenzab57@hotmail.com

Jane Burt is a Senior Consultant with Emerald Network with a background in environmental learning and socio-ecological justice and believes that education and learning can be designed as political acts with transformative force. She has worked in the theatre of the oppressed in a water justice context, which led to transformative learning for environmental justice in social movements and governance structures. She became involved in evaluation as a learning expert, weaving together evaluation and transformative learning. jane.emeraldnetwork@gmail.com

Eko Ruddy Cahyadi is an Assistant Professor of management at the Institut Pertanian Bogor (IPB), Indonesia. He has research interests in operation and supply chain management, project management and impact evaluation. He is a consultant for the United Nations Industrial Development Organization (UNIDO) as a monitoring and evaluation expert on the SMART Fish project in Indonesia. He received his PhD in economics from Leibniz University in Hannover, Germany. ekocahya@yahoo.com

John Colvin, Director of Emerald Network Ltd, has more than 35 years of experience working in multilevel, multi-stakeholder systems change through policy, research, evaluation, design, process facilitation and social innovation. He has extensive experience globally and in Africa, China, Europe and South Asia, having worked widely on climate-compatible development, conservation, land use change, food security, river basin planning, sustainable infrastructure and urban planning. He has worked on assignments for multilateral, bilateral and government agencies; foundations; international non-governmental organizations; and universities and has advised multiple organizations on policy, strategy and practice. Over the past 10 years, he has developed a leading-edge evaluation practice informed by complexity and systems perspectives and transformational learning approaches. john.colvin.emeraldnetwork@gmail.com

Johannes Dobinger is Chief of the Independent Evaluation Division at the Office of Evaluation and Internal Oversight at the United Nations Industrial Development Organization (EIO/IED-UNIDO). He has gained his professional experience mostly in Austria and Central and South America, promoting sustainable industrial development. A significant part of his career was dedicated to evaluation, mostly of industrial development projects and programmes around the world. He holds academic degrees in civil engineering and economics and business administration. J.DOBINGER@unido.org

Fabrizio Felloni is Deputy Director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD). He was previously a Lead Evaluation Specialist there and an Evaluation Specialist at the United Nations Development Programme (UNDP). He has led project, country programme and corporate evaluations in Africa, Asia, Latin America and Eastern Europe for about 20 years. He is the author and co-author of articles published in peer-reviewed journals. f.felloni@ifad.org

Hur Hassnain is a design, evaluation and accountability expert. He has some 20 years of experience in designing and conducting evaluations of 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 United Nations agencies, the European Commission, and a variety of public and private donors. He is an expert in experimental and quasi-experimental design and participatory and theory-based evaluation approaches and has developed innovative tools that have been used in contexts of violence and conflict and presented internationally. Hur is the founder of the Pakistan Evaluation Association (PEA) and a two-term elected IDEAS board member. At IDEAS, Hur convenes the Thematic Interest Group on Evaluations in Fragility, Conflict and Violence (EvalFCV-ITIG). hurhassnain@hotmail.com

Girma Earo Kumbi is a Principal Evaluation Officer in Independent Development Evaluation of the African Development Bank. Previously, he worked for the World Bank and the government of Ethiopia. He has more than 19 years of experience in development and evaluation. G.KUMBI@AFDB.ORG

Jessica Kyle is Director of Climate Change and Sustainability at ICF. She has 20 years of experience in global programme evaluation, with a focus on climate finance and development programmes. She has led complex evaluations for the Climate Investment Funds (CIF), the Green Climate Fund (GCF) and the Global Environment Facility (GEF), as well as for World Bank-administered multi-donor trust funds and bilateral development agencies. For the past few years, Jessica has also been engaged with the CIF's Transformational Change Learning Partnership as a participant and contributor. She holds a master's in ecological economics, value, and policy from Rensselaer Polytechnic Institute, United States. Jessica.Kyle@icf.com

Tim Larson is President of Ross Strategic, a U.S.-based consultancy specializing in strategy, evaluation and learning, and collaborative process support related to the environment and development. He has supported programmes and initiatives related to transformational change in the context of climate finance and action for diverse clients, including the Climate Investment Funds (CIF), Children's Investment Fund Foundation (CIFF), ClimateWorks Foundation, the U.S. Environmental Protection Agency (EPA) and the International Labour Organization (ILO). Tim is a lead facilitator supporting the CIF's Transformational Change Learning Partnership. tlarson@rossstrategic.com

Thuy Thu Le is an Evaluation Officer at the United Nations Industrial Development Organization (UNIDO). Before that, she worked for the International Fund for Agricultural Development (IFAD), the Food and Agriculture Organization of the United Nations (FAO), the Asian Development Bank Institute (ADB) and the International Federation of Red Cross and Red Crescent Societies. For the last 20 years, she has conducted and managed evaluations, research studies and impact assessment of development programmes and projects around the world across thematic areas including environmental management, private sector development and poverty reduction. Her research interests include sustainable industrial development, impact evaluation and information communication technologies for rural development. She specializes in results-based management and evaluation. She holds a master's in international development. t.le@unido.org

Cristina Magro has developed her career in linguistics as a full Professor at the Federal University of Minas Gerais, Brazil, as well as an Associate Professor, Visiting Scholar and researcher at academic institutions in Brazil and abroad. As a Senior Education Expert Consultant, Cristina worked at the Ministry of Education in Timor-Leste under contracts run by the World Bank and the United Nations Children's Fund (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. She devoted more than 25 years to research in systems thinking and teaching systems thinking to non-academic audiences. She is the former Secretary-General for IDEAS and a current member of the International Evaluation Academy Council (IEAc). Cristina holds a PhD in language sciences from the State University of Campinas (UNICAMP – Brazil), is an International Program for Development Evaluation Training (IPDET) graduate and is certified in Institutional Assessment by the Brazilian Ministry of Education and Culture. magro.cristina@gmail.com

Sam McPherson, with more than 25 years of experience in international development, is a Partner at Itad – a consultancy providing strategy, monitoring, evaluation and learning services to make international development more effective. At Itad, Sam oversees the health and education portfolio and is responsible for external partnerships. He has programmatic and evaluation expertise in global health architecture, health systems strengthening

evaluation, knowledge management, strategic planning, results-based management assessment and adaptive management. Sam was the team lead for the Transformational Change Learning Partnership Climate Investment Funds (CIF-TCLP) evaluation. Sam.McPherson@itad.com

Jonathan (Jonny) A. Morell is an organizational psychologist with extensive experience in the theory and practice of programme evaluation. His research focus is on how complexity science can be applied to models, methods and metrics as those constructs are applied in the field of evaluation. He is conducting research on how constructs drawn from evolutionary biology and ecology can be used to understand how programmes work and what outcomes they produce. Jonny was awarded the American Evaluation Association's Paul F. Lazarsfeld Evaluation Theory Award. He is Editor-in-Chief Emeritus of *Evaluation and Program Planning* and owner of 4.669... Evaluation and Planning. He holds a PhD in social psychology from Northwestern University, United States. jamorell@jamorell.com

Linda Morra Imas is interim President of a new initiative – the International Evaluation Academy (IEAc). She is the co-founder of the International Program of Development Evaluation Training (IPDET) and has been its Co-Director and a Chief Instructor for 17 years. She was a Chief Evaluation Officer and evaluation capacity-building adviser for the World Bank Group, working in its private and public sectors; before that, she was a Senior Director at the U.S. Government Accountability Office covering evaluations of education and labour programmes for the U.S. Congress. Now a semi-retired independent consultant, she advises and provides training on monitoring and evaluation. Among her many publications, she co-authored the text on development evaluation, *The Road to Results: Designing and Conducting Effective Development Evaluations*, now translated into eight languages. She is an honorary member of IDEAS and is known for her work on development evaluation competencies and professionalization. lindag1000@gmail.com

Chimwemwe A. P. S. Msukwa is a Senior Practitioner of transformational learning with Emerald Network Ltd and Team Leader for Development Technical Assistance Services (DeTAS), a consulting company in Malawi. He has more than 30 years of designing and providing capacity development services aimed at participatory community development approaches and transformation of local governance institutions, mindsets and practices for socioeconomic development. He has conducted evaluations for

transformational learning in food and nutrition, natural resource management, public health, education, water hygiene and sanitation, and governance. His practice covers the Horn of Africa, Madagascar, Malawi, Mozambique, South Africa, Uganda, Zambia and Zimbabwe. He has a PhD in development studies. cmsukwa@yahoo.co.uk

Mutizwa Mukute is the Director of the Zimbabwean Consultancy Social Learning and Innovation Ltd and a Senior Consultant with Emerald Network Ltd. The desire to understand how well a new agro-ecology and community development curriculum in Zimbabwe, implemented by a rare combination of non-governmental organizations, universities and government departments, had been designed and what could be learned from it initially inspired his evaluation interest. Since then, he has conducted evaluations that have grown in complexity, scale and purpose. His strong interest in just, transformational learning and change influenced his PhD study in environmental education, which sought to explore and expand farmer learning in sustainable agriculture workplace contexts of Southern Africa. mmukute@gmail.com

Ada Ocampo is a Peruvian sociologist who has worked for the United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP), the United Nations Population Fund (UNFPA) and the International Fund for Agricultural Development (IFAD) from 2000 to 2020 in different countries of North and Latin America, Africa and Asia. Ada is one of the founders of the International Organization for Cooperation in Evaluation (IOCE) and the Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC) and is currently president of IDEAS. She has a master's degree in planning and development management from the University of Wales, United Kingdom. president@ideas-global.org

Zenda Ofir is a South African evaluator and scientist. Over two decades, she has worked at the intersection of evaluation, development and research, conducting evaluation assignments across areas and sectors and from the local to the global level, with a focus on Africa and Asia. She has a special interest in societal dispositions, the relationship between humanity and nature, and supporting urgently needed large-scale transformations. She advises international organizations and serves on the editorial boards of several evaluation journals. A former President of the African Evaluation Association (AfrEA), Vice President of the International Organisation for

Cooperation in Evaluation (IOCE) and board member of the American Evaluation Association (AEA), she served as Vice President and interim President of IDEAS and a Lead Steward in the SDG Transformations Forum; she is chair of the Council of the International Evaluation Academy (IEAc). She is an Honorary Professor at Stellenbosch University, South Africa, and was a 2019 Fellow at the Robert Bosch Academy in Berlin. In her early career, she was a Senior Programme Manager at a national science council in South Africa, and later the Director of Research of the University of Pretoria. She has a PhD in ecological chemistry. zenda.ofir@gmail.com

Michael Quinn Patton is an independent consultant based in Minnesota and former President of the American Evaluation Association (AEA). He has authored eight major evaluation books, including *Utilization-Focused Evaluation and Qualitative Research and Evaluation Methods*, used in more than 500 universities worldwide. He has also authored books on practical evaluation, creative evaluation and developmental evaluation – applying systems thinking and complexity concepts to enhance innovation and use. In 2018, he published books on principles-focused evaluation and facilitating evaluation. His new book on evaluating global systems transformations is *Blue Marble Evaluation*. mqpatton@prodigy.net

Robert Picciotto is an Adjunct Professor at the University of Auckland and a senior independent evaluation adviser to the New Zealand Ministry of Foreign Affairs and Trade. He retired from the World Bank in 2002 after holding several management positions, including Vice President, Corporate Planning and Budgeting and Director-General of the Independent Evaluation Group. He is a graduate of Princeton University, United States. r.picciotto@jeunot.net

Khaled Rajab is the Vice Chair of the Palestinian Evaluation Association (PEA) and a development and evaluation expert with a wide range of experience in programme design, institutional building, strategic planning, needs assessment and monitoring and evaluation, including work on public and private sector development. Dr. Rajab is an Adjunct Professor in the master's programme in local governance, School of Law and Public Administration, Birzeit University, Palestine. He holds a PhD from the School of Governance at Maastricht University, Netherlands, and a master's degree in environmental engineering from the Johns Hopkins University, United States. khaledwr@gmail.com

Pablo Rodríguez-Bilella is an Argentinian researcher and evaluator with more than 15 years of progressively responsible positions in applied and academic social sciences: evaluation, planning, community participation and programme management. He is experienced in bridging evaluation research and policy through interdisciplinary work and is interested in linking issues of relevance and applicability with high standards of evaluation quality. His skills include monitoring and evaluation, action research, public consultation, participatory rural appraisal, sample surveys, case studies, project design and review and pilot project implementation. Pablo has been a board member of the Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC), the International Organisation for Cooperation in Evaluation (IOCE), EvalPartners, IDEAS and the Argentine Evaluation Network (Red EvaluAR). In the past 10 years, he has been lecturing on evaluation research methods and supervising undergraduate and graduate students in evaluation topics. He is a member of the implementation team of EvalParticipativa, the community of practice on participative evaluation in Latin America and the Caribbean. pablo67@gmail.com

Silvia Salinas Mulder is a Bolivian anthropologist and feminist with a specialization in Andean studies and a master's degree in decentralization and public administration. She has 30 years of experience as senior consultant, researcher and evaluator in the social development sector and is recognized for her specialized albeit systemic and creative approaches to key poverty, development and exclusion topics. She is an advocate for and facilitator of transformative planning, management and evaluation processes and is particularly interested in power issues, ethics and intercultural relations. Between 2016 and June 2018, she was President of the Bolivian Monitoring and Evaluation Network (REDMEBOL). Since December 2017, she has been chair of the Latin American and Caribbean Monitoring, Evaluation and Systematization Network (ReLAC). She was recently elected president of the International Organization for Cooperation in Evaluation (IOCE). Sylvia is an international speaker and social innovation entrepreneur and has written a vast number of publications on evaluation, gender and other development and human rights topics. ssalinasmu@hotmail.es

Gerardo Sánchez-Romero is an evaluation consultant on projects and policies related to inclusive development, working for governments and the public and private sectors in Mexico and Latin America. He is the founder and Director of Mutua S.C., an applied social research consulting firm based in Mexico. He is the co-chair of EvalYouth Latin America and the Caribbean

and a former board member of the National Academy of Evaluation of Mexico (ACEVAL), the voluntary organization for professional evaluation of Mexico. He has a degree in sociology from the National Autonomous University in Mexico (UNAM) and a master's degree in applied cultural analysis from the University of Copenhagen, Denmark. geruzko@gmail.com

Matthew Savage is Director of Oxford Consulting Partners and an international development economist specializing in international climate finance. He has led a number of high-profile evaluations, including those of the Climate Investment Funds (CIF), Energy Sector Management Assistance Program (ESMAP) and the Kigali Cooling Efficiency Program (K-CEP). He supported development of the U.K. International Climate Fund and is currently seconded to the UK 2021 United Nations Climate Change Conference (UK COP26) Presidency. He worked for the International Finance Corporation (IFC) as Senior Operations Manager, leading regional climate change advisory operations. He holds higher degrees from the University of Oxford and the Netherlands Business School. matthew.savage@oxfordconsult.co.uk

Adeline Sibanda is the founder and Managing Director of ADESIM Developments. Adeline has experience in strategy, programme design, planning, monitoring and evaluation, with a specific focus on gender-responsive and human rights-based evaluations. She works throughout Africa and beyond with civil society organizations, community-based organizations, governments, donors and development agents. She was co-chair of EvalPartners from January 2019 to December 2020, past President of the International Organisation for Cooperation in Evaluation (IOCE) from 2018 to April 2020 and past president of the African Evaluation Association (AfrEA) from 2015 to 2019. Adeline has been raising awareness and advocating for the Made In Africa Evaluation, an AfrEA initiative promoting evaluation practice rooted in African contexts and priorities, and the South-South Cooperation in Evaluation, an initiative of five regional evaluation associations from the Global South. She holds a BSc honours degree in politics and administration and a master's in public administration. troparg@yahoo.com

Daniel Svoboda is Head of the Secretariat of the Czech Evaluation Society (CES) and Coordinator and Head of Administration of IDEAS. He specializes in project management and sustainable development. He has been an external monitor of the European Commission for the environmental programme LIFE since 2009. He was coordinator of the European Program for

Development Evaluation Training (EPDET) between 2007 and 2017 and is currently a lecturer of project management and evaluations at the Czech University of Life Sciences. He was a lecturer in the International Program for Development Evaluation Training (IPDET) at the University of Bern in 2018 and 2019, a member of the Reference Group on Evaluations of the Czech Ministry of Foreign Affairs until 2019 and a mentor on evaluations at the Czech Ministry of Health. svoboda@dww.cz

Juha I. Uitto is Director of the Global Environment Facility Independent Evaluation Office (GEF IEO). He has been a full-time evaluator for longer than two decades with the GEF and the United Nations Development Programme (UNDP), conducting and managing many evaluations at the global, regional and national levels, in particular at the nexus of development and the environment. Juha spent the 1990s with the United Nations University as coordinator of environmental and sustainable development research and training programmes. He has worked at the Nordic Africa Institute and the International Fund for Agricultural Development (IFAD) and has held visiting positions at Kyoto University and Rutgers, the State University of New Jersey. Educated at the Universities of Helsinki and Lund, he holds a PhD in social and economic geography and has published widely in areas related to sustainable development, environmental hazards and evaluation. juitto@thegef.org

Rob D. van den Berg is a Visiting Professor in the Department for International Development of King's College, London, and an honorary associate of the Institute of Development Studies in Brighton. He is a member of the Advisory Council of Wilton Park in the United Kingdom and of the Advisory Panel for the Office of Independent Evaluation of the International Fund for Agricultural Development (IFAD). He serves on the Advisory Group on Evaluation and Learning of the Climate Investment Funds and the Steering Committee of the Transformational Change Learning Partnership. From 2014 to 2020, he was President of IDEAS. Before that, his career took him from the Dutch Ministry of Foreign Affairs to the World Bank, where he was Director of the Independent Evaluation Office of the Global Environment Facility (GEF IEO). Rob has published extensively about development, history, research and evaluation and has co-edited seven books. rdwinterberg@gmail.com

Josephine Watera is Head of the Monitoring and Evaluation Division of the Parliament of Uganda. She has 14 years of experience in Parliament in

the areas of performance management, monitoring and evaluation and evidence-based systems. She was a board member and convener of the Government Evaluation IDEAS Thematic Interest Group (Go-EVAL ITIG), an advisory committee member of the Africa Center for Evidence (ACE), University of Johannesburg and Vice President of the Uganda Evaluation Association (UEA). Josephine was a member of the steering committee of the Evaluation for Transformation Change Award presented at the IDEAS Global Assembly in the Czech Republic. She was a member of a team reviewing the African Evaluation Guidelines by the Africa Evaluation Association (AfrEA). She has mentored a number of young and emerging evaluators under the EvalYouth International Mentorship Program of EvalPartners. She received an award as the most promising evaluator from a developing country by the American Evaluation Association (AEA) in 2015. She is a PhD candidate at the University of Cape Town and a triple master's degree holder in monitoring and evaluation, business administration, and social sector planning and management. jowatera@gmail.com

Sonal Zaveri, based in India, has more than 30 years of experience in strategic planning, programme design, capacity building, mentoring and evaluation. She has worked in more than 25 countries with community-based organizations, non-governmental organizations, international non-governmental organizations, foundations, academic institutions, governments and multilateral organizations. Sonal's interests relate to how rights, participation, transformation and gender are addressed through collaborative, culturally responsive, learning-oriented evaluation approaches and methodologies. She has multiple publications, including book chapters, journal articles, monographs and manuals. Her sectoral experience includes HIV/AIDS, gender, child rights, early childhood development, education, community empowerment, health promotion and prevention, livelihood and disability. She holds regional and global leadership positions on many boards and communities of evaluators such as co-chair EvalGender+, Coordinator of Gender and Equity Network South Asia, GENSA and Founder Board Member of Community of Evaluators South Asia. She has a PhD in social work and was a Fleishman Fellow at Duke University, United States. sonalzaveri@gmail.com

Aaron E. Zazueta is an evaluator specializing in complex transformational processes and sustainable development policies and programmes. In the last 22 years, he has evaluated programmes and policies of the Global Environment Facility (GEF), the Green Climate Fund (GCF), the World Bank, the United Nations Industrial Development Organization (UNIDO); the

Food and Agriculture Organization of the United Nations (FAO), the U.S. Agency for International Development (USAID), the International Development Bank (IDB) and other bilateral institutions and private foundations. He was Chief Evaluation Officer and team leader on impact evaluation for the GEF Independent Evaluation Office. He was Regional Director for Latin America at the Center for International Development and Environment at the World Resources Institute, where he worked with governments, non-governmental organizations and community groups on participatory policy making in Latin America. During the 1980s, he supervised implementation community development programmes in Latin America. He has authored publications on environmental policy, citizen participation, sustainable development and evaluation. He holds a PhD in anthropology from the University of California, United States. a@zazueta.com

IDEAS Books

1. *Development Evaluation in Times of Turbulence: Dealing with Crises That Endanger Our Future*. Ray C. Rist, Marie-Helene Boily and Frederic Martin, eds. Washington, DC: World Bank, 2013. <https://tinyurl.com/4hwzhezv>
2. *Development Evaluation in Times of Turbulence: Dealing with Crises That Endanger Our Future*. Ray C. Rist, Marie-Helene Boily and Frederic Martin, eds. Washington, DC: World Bank, 2013. <https://tinyurl.com/4hwzhezv>
3. *Poverty, Inequality, and Evaluation: Changing Perspectives*. Ray C. Rist, Marie-Helene Boily and Frederic Martin, eds. Washington, DC: World Bank, 2015. <https://tinyurl.com/ru8xewyx>
4. *Evaluation for Agenda 2030: Providing Evidence for Progress and Sustainability*. Rob D. van den Berg, Indran Naidoo and Susan D. Tamondong, eds. Exeter, UK: IDEAS, 2017. <https://ideas-global.org/evaluation-for-agenda-2030/>
5. *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*. Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder, eds. Exeter, UK: IDEAS, 2019. <https://ideas-global.org/transformational-evaluation/>
6. *Evaluación para el cambio transformacional. Oportunidades y desafíos para los Objetivos de Desarrollo Sostenible*. Rob D. van den Berg, Cristina Magro and Silvia Salinas Mulder, eds. Exeter, UK: IDEAS. 2021.
7. *Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners*. Hur Hassnain, Lauren Kelly and Simona Somma, eds. Exeter, UK: IDEAS, 2021.
8. *Transformational Evaluation for the Global Crises of Our Times*. Rob D. van den Berg, Cristina Magro and Marie-Hélène Adrien, eds. Exeter, UK: IDEAS, 2021.

The COVID-19 pandemic has demonstrated the enormous challenges humanity is facing. It has been facilitated by other crises as climate change, biodiversity loss, economic exploitation, and increased inequity and inequality. The UN Agenda 2030 and the Paris Agreement on climate change call for transformational change of our societies, our economies and our interaction with the environment. Evaluation is tasked to bring rigorous evidence to support transformation at all levels, from local to global. This book explores how the future of the evaluation profession can take shape in 18 chapters from authors from all over the world, from North and South, East and West, and from Indigenous and Decolonized voices to integrative perspectives for a truly sustainable future. It builds on what was discussed at the IDEAS Global Assembly in October 2019 in Prague and follows through by opening trajectories towards supporting transformation aimed at solving the global crises of our times.

By combining practical experiences with perspectives drawn from new initiatives, this book offers invaluable insights into how evaluation can be transformed to support transformational change on the global stage.

Indran A. Naidoo, Director of the Office of Independent Evaluation of IFAD

Across continents, educational systems, and historical complexities, this book builds up the language we all should speak about our field. A mandatory read for all young evaluators.

Weronika Felcis, Board member of EES and Secretary of IOCE

After reading these chapters you will have a sharper look at what is relevant when managing or doing an evaluation, and you will notice that 'business as usual' will no longer be an option.

Janett Salvador, Co-founder of ACEVAL, Former Treasurer of ReLAC

This book offers original, visionary discourse and critical perspectives on the challenges evaluation is facing in the post COVID-19 pandemic era.

Doha Abdelhamid, Member of the Egyptian Academy of Scientific Research and Technology

Published by: IDEAS, 2021
ISBN (paper): 978-1-9168982-0-2
ISBN (electronic): 978-1-9999329-9-2

