

Towards Sustainable Futures

The Role of Evaluation

**Edited by Ida Lindkvist, Per Øyvind Bastøe
and Kim Forss**

First published 2024

ISBN: 978-1-032-71488-2 (hbk)

ISBN: 978-1-032-72287-0 (pbk)

ISBN: 978-1-032-71493-6 (ebk)

1 Introduction

Ida Lindkvist, Kim Forss and Per Øyvind Bastøe

(CC-BY-NC-ND) 4.0

DOI: 10.4324/9781032714936-1

The funder of the Open Access version of this chapter is
International Development Evaluation Association.



ROUTLEDGE

Routledge

Taylor & Francis Group

NEW YORK AND LONDON

1 Introduction

Ida Lindkvist, Kim Forss and Per Øyvind Bastøe

Background and purpose

If it cannot be sustained, is it worth doing? Sometimes the answer to this question is a roaring yes. Single bouts of action resulting in saved lives or preventing one-time catastrophes are clearly worthwhile. Other times the answer is no; we are only interested in projects, policies or programmes if they can be sustained and as long as they do no harm to our planet and to us who live here. Global warming is occurring at an unprecedented rate. UN Secretary-General António Guterres called the 2021 report from the Intergovernmental Panel on Climate Change (IPCC, 2021) a ‘code red for humanity’¹. Threats to biodiversity mean that a majority of land surface is altered with ‘ecosystems and biodiversity showing rapid decline’, ‘more species with global extinction now than ever before’ (IPBES, 2019, p. xv) and ‘globally, local varieties and breeds of domesticated plants and animals are disappearing. This loss of diversity, including genetic diversity, poses a serious risk to global food security by undermining the resilience of many agricultural systems to threats such as pests, pathogens and climate change’ (IPBES, 2019, p. xvi).

Making sure that our practices remain or are changed in ways that are sustainable is arguably the most important challenge of our time. Globally, sustainability is acknowledged as one of six key areas to consider in evaluation; however, it is often treated as an afterthought in evaluation terms of references. It is even used very differently. Sometimes sustainability in evaluation refers to preservation of worthwhile practices; other times it refers to change to avoid destruction. There is also a strong normative aspect of sustainability. What is deemed sustainable and worthwhile preserving, or in need of change, differs. Some refer to single projects or policies that need to change or need to be preserved; others like the Sustainable Development Goals link sustainability to specific goals and targets that create a roadmap of what sustainable development looks like for the entire globe. Some talk about local change; others argue that global and systemic change is required to ensure sustainability. Even the time aspect differs from a couple of years to millennia when we discuss sustainable practices.

Sustainability as a concept in evaluation is *essentially a contested concept* (Gallie, 1956).² The concept holds different meanings to different groups of people, it cannot be pinned down descriptively, and it often contains a strong normative element. Collier et al. argue that this gives rise to *a major source of difficulty in both theory and empirical analysis. . . . This in turn influences the coherence of research and the cumulation of findings in the study of politics* (2006, pp. 211–212). When we add the methodological challenges related to the time dimension, it is no wonder that it is tempting for evaluators to respond to sustainability questions with a brief speculative section concluding with perhaps and time will tell. This is particularly true if commissioners are more interested in immediate effects here and now, and sustainability is discussed more as an afterthought.

As we argue that sustainability is a contested concept, there is no and will be no agreement with respect to what sustainability entails in society, and thereby nor in evaluation. Due to the essential contestedness of the concept, various groups will stress different facets of sustainability as more important, and evaluators should be aware that these groups could be in conflict. Unless we acknowledge this and tackle it upfront, we will fail to learn from each other and risk spending our time in the trenches rather than ensuring that our societies change when required and are able to preserve sustainable practices when these are identified. Furthermore, due to the strong normative aspect of sustainability, research and objective evaluation will only take us so far. Concepts of justice, meaningful dialogue and fairness will also need to be considered.

The purpose of this book is to explore the role of evaluation in moving towards sustainable futures. Evaluation has a potential to be an essential tool for understanding and addressing sustainability. It is sometimes said that *if you are not part of the solution, you are probably part of the problem*. So how can and should evaluation become properly and significantly part of the solution in respect of the ‘red alert’? This is what we purport to discuss in this book. We address ourselves to the evaluation community generally, to all who do evaluations, who commission evaluations, use evaluations and have an interest in the topic.

Throughout this book, the fact that sustainability is a contested concept will show up – again and again. It may be quite frustrating for authors as well as for readers. However, simplification when issues are complex tends to be more harmful than helpful. In this book we attempt to portray some of the many ways sustainability can be understood. We refrain from proposing or recommending a single definition or methodological approach; however, we argue that it is essential that the heterogeneity in the different meanings and understandings come to light as this is essential to improve the quality and relevance of evaluation of sustainability.

Sustainability as ‘an essentially contested concept’

So, what does it mean for evaluation, that sustainability is a contested concept?

Gallie argues that ‘essentially contested concepts’ inevitably involve disputes about their proper uses on the part of its users’ (Collier et al., 2006). They maintain that contested concepts are characterised by the following:³ (I) they involve a valued achievement; (II) explanations of worth include reference to different features (diverse describability), however they all recognise a basic way to understand sustainability (original exemplar) and recognise that it is possible to discuss sustainability in this way (reciprocal recognition); (III) there is progressive competition; (IV) the achievement is ‘open in character’ and (V) the achievement is internally complex.

Sustainability as a valued achievement

An essentially contested concept signifies a valued achievement. It is something that we should aspire to. The aspirational nature of the concept can be found in some of its usages among international organisations, companies and governments: The General Assembly of the UN adopted the ‘Sustainable’ development goals in 2015 as universal goals the world should aspire to.

The Dow Jones has its own ‘Sustainability’ World Index, and many of the world’s largest companies report on sustainability and mention it as a leadership value. Amazon writes: ‘At Amazon, we’re committed to and invested in sustainability because it’s a win all around – it’s good for the planet, for business, for our customers, and for our communities’ (Amazon, 2022). Morningstar, an American financial services firm, considered powerful for its investment advice has developed a ‘sustainability’ rating – which is used to assess the extent of risk of companies or portfolios in terms of being not ‘sustainable’. While the accuracy of the reporting can be discussed, being considered sustainable can have effects on stock value and long-term survival.

In 2022, the European Union adopted the EU Corporate ‘Sustainability’ reporting directive (EU, 2022). Countries are also submitting Voluntary National Reviews on sustainability, where they report on progress towards Sustainable Development Goals. In the examples above, sustainability and being sustainable refers to something positive. However, it is also clear that the positive connotation of the word sustainability/sustainable is linked to what is to be sustained. We should be able to take care of our planet, as we do not have a second; we should ensure that poverty reduction interventions are sustainable and so forth.

The international evaluation criterion; ‘sustainability’ – does not encompass all aspects of sustainability, although it does include reference to ‘financial, economic, social, environmental, and institutional capacities of the systems’ (OECD, 2021, p. 71). This was done to ensure universal use; however, given the different emphasis different groups put on different facets it has also made the criterion open to criticism. We will return to this in the discussion. For now, evaluators interested in evaluating sustainability need to understand not only

what constitutes a valued achievement in the evaluation but also how whatever is being evaluated affects other achievements that might be deemed valuable.

Diverse descriptibility, original exemplars and reciprocal recognition

While almost everyone will agree that sustainability is good, different actors describe very different processes when explaining what is required to get to ‘sustainable’. In Collier et al. (2006, p. 217) they explain that the diverse descriptibility of contested concepts *may involve an exclusive emphasis on one or another facet of the concept. Alternatively, as Gallie points out, different facets may be emphasized to various degrees.*

Yet, while sustainability can be explained and achieved in very different ways, it is nevertheless anchored in an original exemplar – or as described by Collier et al., *‘the contested concept is seen as anchored in an original exemplar whose “authority is acknowledged by all the contestant users”’* (Gallie 1956, discussed in Collier et al. 2006, p. 219). The OECD International evaluation criterion ‘sustainability’ is defined as: ‘Will the benefits last?; The extent to which the net benefits of the intervention continue or are likely to continue’ (OECD, 2021, p. 71). A dictionary definition of sustainability reads: *‘the quality of being able to continue over a period of time:⁴’*; another; *‘of, relating to, or being a method of harvesting or using a resource so that the resource is not depleted or permanently damaged’⁵.*

Yet, while most will recognise this original understanding of sustainability, or core content if you will, different groups emphasise different types of sustainability and have widely different time perspectives.

Environmental sustainability often refers to climate change and acting within our planetary boundaries. If we talk about environmental sustainability, then no action can be sustainable if the long-term consequence is that we harm the planet we inhabit. The difference between ‘one way or another’ is the difference between having to change our ways due to ecological collapse/uncontrolled global warming or reducing global emissions and taking care of the environment now.

Others subscribe to a broader meaning of sustainability, where sustainability refers to environmental, economic *and* social sustainability together as the most important tasks of our time. The Sustainable Development Goals (SDGs) adopted in 2015 have put these three interlinked aspects of development as key to a sustainable future.

Yet others have a shorter time horizon and are concerned with humanity and our way of life, where the ‘resource’ being depleted could be the ways we want to live, our values and what kind of society we would like to have. Or depletion of financial resources required for the continuation of what we are interested in.

For evaluation this diverse understanding of sustainability clearly has consequences. First, the evaluator should recognise what type of sustainability they are investigating, and, second, that more aspects of sustainability may be applicable

than what the commissioner first envisions. Under specific circumstances different actors may emphasise different parts of sustainability, and different sustainability goals may be incommensurable.

Progressive competition and normative aspects

Gallie writes that ‘competition between scientific hypotheses works successfully largely because there are acknowledged general methods or principles for deciding between rival hypotheses, for all that these methods or principles may never be completely formalised or finally agreed’ (Gallie, 1956, p. 179). For contested concepts, however, there is no process to determine which concept should prevail, because the type of emphasis that is placed on various aspects is partly dependent on where the evaluator or commissioner stands and values. There is no scientific or logical process that can state that one value or one way to view sustainability is inherently better than all others.

Instead, we have endless debates and criticism between different sustainability camps. This is illustrated by the discussion of the international evaluation criterion sustainability. One of the arguments has been that the criteria do not adequately encompass the 2030 agenda narrative and current policy priorities. Consultation found that some criticised the criteria for being project focused and not sufficiently taking the complexity of interventions into account, including trade-offs and the integration of human rights and gender equality (OECD, 2019, p. 3).

Some of these criticisms emerge from a concern about how development traditionally has been approached:

The fragmentation, the reductionist thinking, the soft power agendas, the belief in outdated models and ideologies at the cost of learning from real success stories, the lack of engagement with concepts and the obsession with ‘methods’ and ‘tools’, the lack of innovation in capacity building that spans the interests and capacities of the Global North and Global South – all of this needs our concerted attention.

(Ofir, 2017)

Consequently, the argument goes, the evaluation approach needs to change. Increased uncertainty and risks require a ‘post-normal’ evaluation practice that reflects unpredictability as well as incompleteness, instability and a plurality of perspectives in value determination. The success of such an approach is whether it is

resilient enough to adapt to the ontological realities of complexity, uncertainty, and contradiction in ways other than being methodologically innovative. That resilience is largely an ethical matter, of evaluators taking full

responsibility for the choices they make in framing and bounding the evaluations they conduct with the public.

(Schwandt, 2019, p. 327)

These views were amplified by Michael Quinn Patton in a blog post in 2020: *I believe the truth is that the DAC criteria communicate a conservative, business-as-usual message that is out of touch with the global climate emergency we face.* As an alternative approach, Patton presented the so-called Blue Marble Evaluation. His reasoning was that fundamental systems transformations are needed to address the global emergency brought on by climate change and related global trends, including the COVID-19 pandemic, which, together, pose existential threats to the future of humanity.

Achievement of sustainability is ‘open in character’

Another aspect of contested concepts is that what it means to be sustainable will change depending on time and context. What actions are required to make something last changes because what would be sufficient to reach a desired state in one context is not enough or appropriate at a different time or in a different context. Sustainability is thus an emerging phenomenon, and not only time and context but also scientific advances shape what can be considered sustainable. Policies and practices that were considered sustainable ten years ago may not be so any longer, and as we all know the pace of scientific advances is faster than ever.

What this openness requires of evaluators is firstly humility; we should all be aware that practices deemed to be sustainable in the present may not be deemed in a positive light in the future, or they may not be sustainable in all contexts. This also means that evaluators need to acknowledge uncertainties and discuss how different future trajectories may make something that looks sustainable in the present unsustainable in the future depending on which of these trajectories. While future studies is a subject that is gaining traction, it could also help to look into the past as is argued by Forss et al. (2020) in *Long-term Perspectives in Evaluation*.

Complexity

According to Gallie, *This achievement must be of an internally complex character, for all that its worth is attributed as a whole* (1956, pp. 171–172). A concept is not necessarily contested because it is complex, but it is more likely that something complex will also be contested. Sustainability is complex for several reasons. First, the concept itself has many and interrelated dimensions; the economic, social and environmental being just one. Second, we are interested in how sustainability is achieved (and lost) and hence in causal mechanisms. Complex causal links are likely to be non-linear, multiple, unstable, and multidirectional.

Third, when turning to the future we must consider risks, uncertainties and multiple, interwoven trajectories of change.

A fourth aspect concerning the nexus of evaluation, complexity and sustainability is whether the evaluation process itself is complex/has to be complex. We do not believe it necessary to discuss whether achievement of sustainability requires a complex process – we simply argue that it does. But the consequence for evaluation of sustainability merits more discussion. The book *Evaluating the Complex: Attribution, Contribution and Beyond* (Forss et al., 2011) argues that evaluating complexity and thereby sustainability requires various methods as evaluations need to be concrete, specific, flexible and innovative – and hence no single method holds the key to all questions. The complexity of a process required to achieve sustainability can lead to surprising and unexpected outcomes, which means that evaluators will need methods to deal with this.

Evaluation of sustainability is not new, but can it be trusted?

Everyone will agree that conclusions on sustainability must be trustworthy. But quality of evaluation is not always up to par. Evaluation quality has been discussed in other Inteval books, see for example *Quality Matters* (Schwartz & Mayne, 2005). One of the ways to approach quality is to check how well evaluations respond to the six international evaluation criteria. Several studies of evaluation quality indicate that the criterion of sustainability is more problematic than any of the other criteria. Not only is it not so common that evaluations conclude on ‘worth or merit’ in respect of this criterion; in addition, when they do, these conclusions are neither valid nor reliable (EBA, 2017).

If we imagine a specific intervention with an objective of achieving a specific development goal, a serious assessment of sustainability does need to consider not only whether the benefits will last into the future, which can be hard enough, but also how the intervention will affect any of the other goals. With a complex view of sustainability, it is not enough to demonstrate that, for example, the health benefits of a vaccine project improve health; the evaluator must also ensure that the factories that produce vaccines do not harm the environment. This can clearly be a daunting task for an evaluator as the evaluation of sustainability may appear unrealistic for any project. Not only does an evaluator need to tackle saying something meaningful about the future – they also need to say something meaningful about a lot of different development objectives the project may not have intended to target.

If we want to improve the assessment of sustainability, it will be necessary to know why the quality of these conclusions is low. Three explanations come to mind:

1. *‘Crowded’ terms of reference.* Many who initiate evaluations have many questions they seek answers to. The six international evaluation criteria are guidelines, and often they are all present to guide evaluation teams, equally often

elaborated in detail. In the conduct of an inquiry, the amount of time that can be devoted to any question is limited. While the questions of relevance, efficiency, effectiveness, impact and coherence often can be answered by similar data and related processes of analysis, sustainability is of a different nature.

2. *The methodological debates* in the evaluation community, as well as the contents of many courses, textbooks and manuals, deal with the most pressing issues, establishing causal links, understanding the intervention and its context, documenting lessons learned and developing recommendations. To analyse sustainability takes the task to the future, and this may require other forms of data, other processes of analysis, and new stakeholders' presence.
3. *Changing from focusing on present concerns to exploring the future and learning from the far past.* In a rough characterisation, it might be said that most of the intellectual effort in an evaluation is spent to understand concerns of the present and to study recent past events. Evaluators focus on the decision needs of today, which are often geared towards budget cycles, where five years is a long period – But to answer the questions around sustainability, evaluators need to understand long-term change and look into the future – often far into the future. The methods and the tools to do so are developed in Futures Research, an interdisciplinary branch with its own journals, its own academic community and its own processes of outreach to decision-makers. It is rare to see evaluators mix with this community, rare to see evaluation proposals build on sophisticated methods to study future sustainability.

Content of the book

Is yet another text on sustainability and evaluation necessary? We have asked ourselves that question and our affirmative response is based on three factors: (1) there is an urgent need for transformational change moving our societies onto trajectories of sustainable change; (2) the task of evaluating sustainability is in an altogether different league of difficulty compared to traditional evaluations of goal achievement, impact, efficiency and effectiveness because of the contested nature of the concept of sustainability; and (3) so far the evaluation community at large has not contributed much; the quality of methods and approaches and hence validity and reliability of conclusions leave much to be desired. This book will not solve all questions, but it does present insights based on current theoretical developments and insightful practice.

Overview of the chapters in the book

As sustainability is a contested concept, there is no, and probably will never be, complete agreement either about what sustainability entails or with respect to how it should be evaluated. The first part of the book introduces the reader

to key debates and challenges related to evaluation of sustainability, while the second part is meant to be inspirational providing examples of methods and applications.

Part 1: From projects to systems and politics

Kim Forss explores why and how change becomes sustainable – or could be found *not* to be sustainable. The chapter raises the question of how resilience can be built into an organisational system to make sure that the achievements of policy objectives are not eroded. Whereas the first question for an evaluation might be termed ‘is this sustainable?’, the next and more important question is ‘how can this be sustained?’ Such questions connect evaluation to management and also point to the importance of monitoring to provide ongoing information on sustainability. The chapter also turns to the imaginative risk management and raises questions concerning appropriate methods to identify threats to sustainability early to mitigate or neutralise their impact.

Ida Lindkvist discusses evaluation of sustainability in the presence of goal conflict. To understand whether projects, programmes or policies are sustainable in the presence of goal conflict, evaluation needs to shed light on the worth and merit of these projects when goals collide. The chapter is focused on conflicts that can arise from green energy projects that can have negative effects on the environment and for indigenous people living in these areas. The main argument is that when goals are incommensurable, that is, when no common scale by which goals can be compared exists, evaluation cannot conclude whether the project is beneficial. Instead, evaluation can help shed light on which values are at stake.

In the third chapter, **Rob D. van den Berg** relies on a large body of literature to discuss the evolution of the concept of sustainability over the past 70 years up to today’s Agenda 2030 with the Sustainable Development Goals. Van den Berg distinguishes three concepts that are currently being used. The oldest is sustainable economic growth, followed by social and economic sustainability, while the most recent sustainability paradigm concerns an adaptive balance between the social, economic and environmental domains. He argues that sustainability is not to be found in interventions but in the systems where interventions take place or which they aim to influence. Evaluation of sustainability needs to take a systems perspective. If we continue to focus on interventions only, evaluation will not be able to contribute to a more sustainable future. The chapter ends with some practical suggestions – for how we can move into sustainability evaluation or a sustainable evaluation.

Saeed Parto’s chapter explores whether and how the gaps between intended objectives under sustainable development and the actual outcomes may be narrowed, using an institutionally informed policy analysis framework. Institutional analysis to inform policy, the chapter argues, should begin with the specification

of the institutions that characterise the context in which change is being sought. The central question for the policymaker and the policy analyst is not how ‘good’ a policy or a mode of governance is in the abstract but how closely the policy resonates with, and can change, the pre-existing conditions and the institutions through which a policy arena is governed.

While most of the chapters of this book focus on sustainability, the chapter of **Mathilde de Goër de Herve** tackles the lack of sustainability, namely disasters, and how the management of such events can hinder or promote sustainable development. The risk of disasters is ever present and is a key challenge that directly affects the well-being of the planet’s inhabitants. Yet, their management can present conflicting results in terms of the contribution to sustainable development, especially given the various possible understandings of what sustainable development is. This is illustrated in the chapter through two mechanisms: first, the transfers of disaster risks, when reducing one risk increases another one, be it the same or a different hazard, in the same or a different place, today or in the future, and for the same or different actors; and second, path dependency, when future disaster risks and their governance are affected by what is happening today.

Given the diverse views of what sustainability entails, it is crucial to rely on different perspectives and lessons from social science and the humanities.

Part 2 Methods and applications

Part 2 provides advice for evaluating sustainability in the private and social sectors, inspiration, and lessons from the social sciences on how to evaluate change, as well as providing an overview over different ways to view sustainability.

In his chapter, **Jens Andersson** argues that circular business models are seen by many as an attractive way for companies to develop new business opportunities, while simultaneously addressing growing social and environmental challenges beyond traditional Corporate Social Responsibility (CSR) programmes. The chapter introduces the circular business model literature with the aim to explore novel approaches to evaluate sustainability. The chapter concludes that the evaluation of sustainability could be further enriched by shifting from a linear to a circular mindset, applying business case thinking to assess the economic dimension of sustainability, and engaging with the quickly evolving sustainability legislation.

While the private sector has mainly been concerned with financial sustainability for long, the incorporation of reporting on other types of sustainability is a more recent development and is still in its infancy. The next chapters discuss both the utility and credibility, but also how we can be aided by artificial intelligence when identifying myriad of reporting.

Per Øyvind Bastøe and Paul Wade discuss in their chapter criteria for determining the credibility and utility of corporate sustainability reporting,

which is now the main tool for a company to voluntarily communicate its performance and impact in environmental, social and governance matters. The assessment framework, with five criteria, is designed for quick screening of whether companies' sustainability reporting is credible. While it does give a comprehensive picture of a company's current state of sustainability, it is not intended to drill deep down to give a fully exhaustive picture of all aspects of a company related to sustainability and over a longer time span. The framework is tested on three selected companies with headquarters in Norway: Hydro, Yara and Wilhelmsen.

Francesco Mazzeo Rinaldi, Giovanni Giuffrida, Salvo Nicotra and Flora Dispinseri explore and reflect on the possibility of using a classification algorithm to analyse and assess the communication profile of companies and organisations engaged in the SDGs. The chapter seeks to answer the following main questions: How to analyse whether companies include SDGs in their strategies through classification algorithms? How to assess the distribution of their commitment across SDGs? Can these algorithms be used to suggest practical actions to improve alignment with the stated goals, making them conform to the expectations of customers and shareholders? And, finally, how to assess the finalisation of governance against sustainability goals using these algorithms? Testing shows that the use of information technologies, specifically of an algorithm capable of classifying textual documents using the 17 UN SDGs as a filter, can help redirect a company's communication and projects, allowing them to become more environmentally friendly.

Kjeld Høgsbro and Olaf Rieper focus on social sustainability and discuss the meaning of sustainability in social services for citizens with complex disabilities as well as guidelines for evaluating social care in a sustainability perspective. Integrated care for citizens with complex disabilities is a challenge for both local and regional authorities, in terms of both quality and financial viability. They explore how incorporating sustainability evaluation into integrated care itself can contribute to more sustainable services. They argue that this will require an evaluation system that is self-sustainable and requires an evaluation system that should be continuously running and include stakeholders from the very bottom where you have services that help citizens up to the municipality level and even higher where they organise these services.

Peter van der Knaap explores the combination of 'sustainability' and 'coherence' in connection to theory-based evaluation and Theories-of-Change (ToCs). The chapter shows how this was done in a case study of efforts to arrive at 'sustainable road safety': a road traffic system that is organised to systematically trigger – or enforce – safe behaviour of road users. The case study provides lessons for policymakers and evaluators – demonstrating alternative evaluation methods that allow for systemic complexity and the prerequisites for transformational change that are indispensable in this field.

It is rational and logical to expect that progress towards sustainable development presupposes an understanding of the mechanisms that shape how and why societies change. To do so, **Tove Heggli Sagmo and Anita Haslie** present social practice theory and discuss how that can be used in evaluation of (and for) sustainable development. The social practice perspective shifts the focus from individual behaviour, attitudes and values towards social practices understood as routines and norms that are shaped in relation to context. The chapter builds on experiences from development evaluation and synthesises a conceptual framework, concretely providing tools for how to integrate the perspective of social practice in evaluations of sustainability.

Alison Pollard focuses on social sustainability. She considers how operationalising the concepts of social capital and well-being and applying human rights-based approaches and participatory techniques can support evaluators in studying social sustainability. To explore how they can aid evaluation of sustainability, she uses two case studies: (1) an evaluation of the sustainability of urban housing demolition and development programme in Shanghai, China, and (2) a final evaluation of the land and rural development project in South Africa. Pollard discusses sustainability in terms of the achievement of social goals such as positive human relationships and social justice, and how the achievement of these goals is expected to evolve over time. Pollard argues that evaluations should consider what social sustainability means to people in a specific context based on their lived experiences and values. The use of participatory techniques and rights-based language enabled the studies to provide insightful findings about social sustainability.

Sanjeev Sridharan, Debra Torok, Abhijit Das, Satish Kumar Singh, Amanda Pereira and April Nakaima argue that as we implement and evaluate interventions that promote sustainable development, we need to focus on *sustainable* impacts and consequently plan for sustainability. Furthermore, given the centrality of reducing inequities in the Sustainable Development Goals, we need to reflect on how evaluative thinking can help sustain interventions that reduce inequities. These ideas are explored through the example of a sustainability evaluation of an intervention focused on addressing Gender Equality in India. Key implications of this chapter include the explicit need to pay attention to timelines of impact, clarification on what is a good enough theory, recognition of initial incompleteness in knowledge and models of learning, evaluative thinking to bring clarity on the role of funders post-funding and evaluation's role in helping understand client needs over time.

Tom Ling and Ananda S. Millard, in their chapter, unpack some of the dimensions of sustainability with the aim of providing clarity and focus for future evaluations. The chapter proposes four levels, which could be described as 'sustainable' and for which there are different evaluation frames. They argue that this approach is helpful in organising how evaluators might conceptualise and analyse levels of outcomes. The model is not intended to serve as a straitjacket for how sustainability should be measured. Rather it should facilitate a more

systematic assessment of sustainability and help evaluators navigate different sustainability dimensions.

Part 3 Conclusions

The final and concluding chapter is a call for action and a philosophical exercise. Taking sustainability more seriously can help preserve the kind of society we want to have for people and the planet. The crux of the argument is that we cannot take the present for granted and assume that things will stay the same. But what roles can evaluation play? The chapter builds on the 15 previous chapters and outlines six specific and concrete roles for evaluation as a practice and as a discipline could play in shaping sustainable futures (the plural indicating that there is not one solution but many depending on context, time, localities and people).

Raising the stakes – real challenges

When the international evaluation criteria were developed, it was assumed that the impact of an intervention should be sustained. If not, the achievement of objectives, the impact created, and the relevance would have been wasted. This has an obvious management logic to it and seems evident by common sense. However, the current international focus on the 17 Sustainable Development Goals poses new challenges. It is no longer the question if the results of an intervention can be sustained. There can hardly be any doubt that the overall processes of change in the world today are not sustainable and need to be radically transformed. The chapters in this book describe evaluation in different sectors and different contexts, and it is not surprising that an overview must emphasise diversity. Nevertheless, there is a series of issues that come across as common experiences and concerns:

1. *Incorporating a systems perspective*: The criteria definition notes the assessment of sustainability includes an examination of the financial, economic, social, environmental and institutional capacities of systems. To this might be added specific questions of culture, gender and poverty that are more focused than these general terms. To understand how something will develop into the future, we need to understand the system which what is evaluated operates within.
2. *Intersectoral questions and goal conflicts*: The evaluation criteria of sustainability focus on whether the net benefits of an intervention will last, acknowledging that potential negative effects might arise in the pursuance of an objective. The Sustainable Development Goals provide a net of interconnected goals, which may be used to identify what is harmed in the pursuance of one goal. The challenge for evaluation here is being able to identify

negative or positive effects on other goals and having the tools to identify goal conflicts and to assess whether the positive effects of one goal might offset the negative effects of others.

3. *Sustainability cannot be understood as a static property.* It is an attribute that would need to change in nature over time. The properties that make something sustainable today may not do so ten years from now. Risk, uncertainties and resilience need to be built into the assessment.
4. *Complexity needs to be recognised and embraced.* Much has been written about evaluation and complexity, also by the Inteval group. The focus is often on how to reduce and manage complexity, at times on avoiding complex questions. When complexity is a fact of life that cannot be neglected, an altogether different approach is needed.

Lessons for evaluation

So far, we have argued that sustainability is a contested subject and that this has several consequences for evaluation. As a contested concept, there is no and will be no agreement with respect to what sustainability entails in society, and thereby nor in evaluation. This holds true even though sustainability is one of the most important challenges of our time. Due to the essential contestedness of the concept, various groups will stress different facets of sustainability as more important, and evaluators should be aware that these groups may be in conflict with each other.

This means that evaluators need to understand not only what commissioners mean by sustainability but also what other groups in society could mean by sustainability in that context and at that time. When different sustainability facets come into conflict, this will require meaningful stakeholder participation, as well as high-quality analysis and application of methods to avoid elite capture. Participatory evaluation is often an excellent tool in project evaluations, but in policy evaluation simple stakeholder participation may not be enough. An altogether different level of political legitimacy, political decisions-making and accountability is called for.

In addition, regardless of what facet of sustainability is emphasised, achievement of sustainability is likely to depend on a complex process where long-term change may be difficult to predict. This requires methods that can deal with non-linear causal relations, interdependency, and system's relationships, and perhaps even more important a healthy dose of humility. Furthermore, what is deemed to be sustainable will change depending on time and context. What is deemed to be sustainable today may be viewed as unsustainable tomorrow. We have argued that evaluators may need to be careful with their claims and open about assumptions that need to hold true for interventions to last.

Finally, the near-universal agreement on sustainability as a valued achievement means that evaluators need to consider the question; Should this last? Is

there something better? The positive connotations of sustainability may bring a bias against negative questions. Evaluators should consider not only whether the benefits will last but also whether they should.

Notes

- 1 Guterres: The IPCC Report is a code red for humanity (unric.org).
- 2 Gallie's article 'Essentially Contested Concepts' struck a nerve when published in (1956) as it described types of concepts that all know and relate to but where it is not possible to find a conceptual strict definition/understanding that will work for all people at all times. Examples include fairness, democracy, rule of law and art.
- 3 Gallie lists the criteria in the following manner: (I): Valued achievement, (II) The achievement is internally complex, (III) Diverse describability, (IV) The achievement is open in character, (V) Reciprocal recognition, (VI) Original exemplar whose authority is recognized by everyone, and (VII) Progressive competition. For reader-friendliness we have altered this list, and grouped (III), (V) and (VII) together.
- 4 SUSTAINABILITY | meaning, definition in Cambridge English Dictionary.
- 5 Sustainable Definition & Meaning – Merriam-Webster.

References

- Amazon. (2022). *Delivering Progress Every Day*. Retrieved from <https://sustainability.aboutamazon.com/>
- Collier, D., Daniel Hidalgo, F., & Olivia Maciuceanu, A. (2006). Essentially contested concepts: Debates and applications. *Journal of Political Ideologies*, 11(3), 211–246.
- EBA. (2017). *Livslängd och livskraft: Vad säger utvärderingar om svenska biståndsinsatsers hållbarhet?*, EBA Rapport 2017:12, Expertgruppen för biståndsanalys.
- European Union. (2022). *Directive (EU) 2022/2464 of the European Parliament and of the Council as Regards Corporate Sustainability Reporting*. Brussels. Official Journal of the European Union. L 322/15.
- Forss, K., Lindkvist, I., & McGillivray, M. (2020). *Long Term Perspectives in Evaluation: Increasing Relevance and Utility*. Oxford and New York: Routledge.
- Forss, K., Marra, M., & Schwartz, R. (2011). *Evaluating the Complex: Attribution, Contribution, and Beyond* (Vol. 1). New Jersey: Transaction Publishers.
- Gallie, W. B. (1956). Essentially Contested Concepts. *Proceedings of the Aristotelian Society*, 56, 167–198. <http://www.jstor.org/stable/4544562>
- IPBES. (2019). Summary for policymakers. In E. S. Brondízio, J. Settele, S. Díaz, H. T. Ngo (Eds.). *Global assessment report of the Intergovernmental science-policy platform on biodiversity and ecosystem services*. Bonn, Germany: IPBES Secretariat.
- IPCC. (2021). Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L., M. I. G. Goldfarb, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, R. Yu, B. Zhou, & O. Yelekçi (Eds.), *Climate change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
- OECD. (2019). *Better Criteria for Better Evaluation: Revised Evaluation Criteria. Definitions and Principles for Use*. Paris, France: OECD Publishing.

- OECD (2021). *Applying Evaluation Criteria Thoughtfully*. Paris, France: OECD Publishing. <https://doi.org/10.1787/543e84ed-en>
- Ofir, Z. (2017, November 14). *The DAC Criteria, Part 11. From Criteria to Design Principles*. Blogpost. November 14, 2017. Retrieved from: <https://zendaofir.com/dac-criteria-part-11/>
- Patton, M. Q. (2020). *Evaluation Criteria for Transformation Webinar Q&A*. Blogpost. March 20, 2020. Retrieved from <https://bluemarbleeval.org/latest/evaluation-criteria-transformation-webinar-qa>
- Schwandt, T. A. (2019). Post-normal evaluation? *Evaluation*, 25(3), 317–329.
- Schwartz, R., & Mayne, J. (2005). *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance Reporting*. New Jersey: Transaction Publishers.