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)

15 Evaluating sustainability – a multidimensional approach

Tom Ling and Ananda S. Millard

DOI: 10.4324/9781032714936-17

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

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

15 Evaluating sustainability – a multidimensional approach

Tom Ling and Ananda S. Millard

What is sustainability, and why is more conceptual clarity important to inform practical evaluation tools?

Sustainability is a good thing in the broadest sense that we urgently need to build a more sustainable relationship between human societies and the biological and ecological systems that make human life possible. But we have also seen that it is challenging to operationalise it as a concept, and even more how to implement it as a practice. Indeed, in many ways it should be viewed as an essentially contested concept (Gallie, 1964). That is to say, while it is widely approved of, in any particular instance, entirely reasonable arguments might lead to very different conclusions. This is because to talk of 'sustainability' immediately raises the question 'sustainability of what and for whom?' We do not want to dissolve these differences, not least because they involve important arguments, but we do want to propose a pragmatic way forward that recognises that sustainability might mean different things in different contexts, at different times in the lifecycle of any type of intervention and may differ from subject area to subject area. Indeed, the dimension of sustainability that might be appropriate at a specific point in time, for a specific subject, and in a particular context may differ vastly from what may be needed in a different instance. In some instances, seeking sustainability of the wrong thing for the time and context may be more damaging than helpful.

In this chapter, we hope to contribute a little order to the debate. This involves recognising that sustainability can be reasonably conceptualised in multiple ways, that clarity about which dimension is being used helps both the evaluation and the practice of sustainability, and that rubrics-based assessment of sustainability can help greatly in bringing clarity and transparency to the evaluation process. Conceptual clarity is important to inform practical solutions to be used in evaluations because without it understanding what individual evaluations mean, and learning from across multiple evaluations, is compromised.

DOI: 10.4324/9781032714936-17

What has been the predominant thinking regarding how sustainability is achieved in the development sector?

The World Commission on Environment and Development (1987) defined sustainable development as development that 'meets the needs of the present without compromising the ability of future generations to meet their own needs'. In 2015 the 2030 Agenda for Sustainable Development continued this emphasis (United Nations, 2015). Since then, there have been various modifications to this approach and not least the introduction of politics as a fourth dimension.[1] But, however, this approach might be developed, at its core, 'sustainability' points to the ability of social world and the biosphere to continue to operate 'without compromising the ability of future generations to meet their own needs'.

This is, by any definition, a grand challenge. It underpins many policies, reports, and commitments and in particular the Sustainable Development Goals. Therefore, a standard question asked of international development initiatives is '(how) does it contribute to sustainability?'. The OECD DAC evaluation criteria[2] have been updated and, for sustainability, concern the extent to which the net benefits of the intervention continue or are likely to continue. The definition notes that assessments should include an examination of the financial, economic, social, environmental and institutional capacities of the systems needed to sustain net benefits over time. This should involve analyses of resilience, risks and potential trade-offs. Depending on the timing of the evaluation, this may require analysing the actual flow of net benefits or estimating the likelihood of net benefits continuing over the medium and long terms.

While the above provides interesting points of departure, this definition does not properly recognise that sustainability is a characteristic of something (and not itself a thing), that it may be evolving over time, and that it can manifest in different ways and in different dimensions. Neither does it sufficiently recognise that there is a dynamic relationship between sustainability of something and the context within which it sits. To understand this dynamic relationship, it is necessary to recognise that as context changes so must the characteristics needed to maintain sustainability.

Equally, there are instances where attaining sustainability may mean the maintaining of sub-standard responses. For example, an expansion of access to education without attention to the quality of the education provided could mean that access to a sub-standard product is sustained without any focus on continued transformation, where access can be a first step and sustained quality can be a second step of a process. Principally, the 'non-harm to future generations' modality does not recognise that sustainability, in most instances, should not be a static point of achievement but rather a dynamic state that should most likely change and adapt to contextual conditions over time and that different interventions should seek to both deliver and measure different dimensions of sustainability. Indeed, in some instances and for some interventions not attaining any form of

sustainability may be preferable. This would apply to efforts which are expected to serve as gateways for interventions that better address known problems (and which should be judged among other things against the criteria of sustainability). The understandable efforts to establish the importance of sustainability in global debates have provided a conceptualisation of sustainability (non-harm to future generations), which is an important rallying cry, but which requires further effort to operationalise it consistently within individual evaluations and evaluation systems more generally. In particular, we propose that it should be understood in a multidimensional way.

Why is examining sustainability in a multidimensional way important and how might this be done?

Evaluating the role of programs and interventions in securing intentional (sustainable) change can be addressed, for example, in relation to delivering system change for greater racial justice in working life evaluators might explore: the numbers of people of colour who can access decent jobs, the availability of information flows about available labour market opportunities for people of colour; incentives for employers to change behaviours (social structures), and changing attitudes and expectations regarding the purpose of work and employment in relation to social justice and productivity (system goals). A satisfactory approach would need to be multidimensional to encompass the various aspects of system change.

Our approach for assessing sustainability is therefore informed by systems thinking, in that we see all dimensions of sustainability as existing within systems, and as multidimensional. It relies on five dimensions and identifies how each dimension exists within a system. These dimensions are in no way hierarchical nor is it suggested that one may be better than the others. In each case, the programme, project or other intervention is seen as a set of actions or an event in a pre-existing complex system, which then interacts with that system to support sustainability (or not).

A rubric-based reflective approach that can help develop practical tools for evaluating sustainability

Here we present the five dimensions of sustainability that underpin our rubric-based approach. This should be regarded as an approach that provides evaluators with a reflective toolkit that can be used to develop specific tools that are catered to the evaluation being conducted. By ‘dimensions’, we mean the differing characteristics of sustainability that can serve to frame what is meant by sustainability in any given case. In some instances, a single intervention will have manifested different dimensions of sustainability over time, in other instances a single dimension will be sufficient to describe the case being evaluated.

Dimensions of sustainability are not necessarily chronological but can be. Also, it is worth noting that different dimensions can relate to each other.

Dimension A – Discovery or gateway activities: the intervention has been developed because there is some reason (based on research, experience or logic) to expect that it should yield positive results, but the evidence for how to achieve results is patchy. The intervention is intended to test and explore one or more approaches which could, if successful, lead to more sustainable intervention or demonstrate that the intervention model is not valuable, or that the intervention model is valuable, but only as a gateway to another intervention. In the latter case the intervention should not be sustainable in any way. Indeed, in these instances the sustainability of the intervention or its effect is a demonstration of failure to progress towards the ultimate goal that underpinned the intervention in the first place.

Strictly speaking, this dimension is one that deliberately does not meet any sustainability criteria. The dimension has been included because we feel that evaluation often overlooks the need to genuinely explore new ways of working and testing interventions. This may contribute to sustainability without itself being (or intending to be) sustainable. Evaluations may be so focused on the importance of sustainability that this important dimension may be missed. Including this dimension can allow evaluators to deliberately highlight the important contribution of discovery to delivering sustainability but without requiring that in an uncertain world sustainability is the only measure of success.

Moreover, this dimension may also serve to highlight that some interventions should by design be exploratory or short-lived and not necessarily sustainable. It should allow us to better understand what might be sustainable, and what should not be sustainable. That the progress such projects achieve is, and should be, short-lived and only contribute to other activities which deserve to be sustainable, or they might be a short-term solution before more effective long-term solutions can be found and/or put in place. For example, the introduction of quotas alone should not be sustainable unless it is accompanied by a whole series of other activities that render quotas meaningless. Essentially making sure that women, and or other minoritised groups, are represented by imposing quotas without addressing the root causes of the unequal representation will not address said root causes. Therefore, ensuring the sustainability of such a system will most likely lead to other problems: such as anger and resentment by non-prioritised parties, and even abuse by groups that are benefited by the quota. In India, for example, efforts to fill local government positions with women means that in certain instances a woman is elected, but her husband or another family member is the one taking up her position. This is an example of an effort that did not target the root causes and hence its sustainability is not beneficial to ensuring equality between men and women.[3]

Dimension B – replication: at the end of the intervention, mechanisms remain in place to ensure that the funding, activities, and outputs that characterised the intervention persist and multiply. This enables the intervention to continue unchanged.

The dimension describes the continuation (persistence) of an intervention. At this level, evaluations would typically review the financial arrangements, skill levels, partnership working, acceptability and contextual characteristics before forming a view on whether, taken as a whole package, the existing modalities were sustainable. For example, responsibility for a sanitation campaign was handed over from an NGO to a local leadership with the skills, capacities and funding to deliver continued sanitation for the target population. This dimension does not expect that the scope of the intervention will be expanded, or modified, and often requires continued financial support. Often this dimension manifests because no mechanism to expand the activity is made or necessary. This level of sustainability may be all that is needed. For example, in the provision of emergency response assistance. However, the same dimension of sustainability may be a demonstration that the service provider has no interest in making the intervention truly sustainable (at a system level, for example).

Dimension C – Programme outcome consolidation: the intended parameters of the intervention are achieved more often and in many places as a result of the experience of the programme (e.g. more people trained, more services provided, more taxes collected) and/or the programme is operationalised in a far broader way than initially.

This dimension can materialise in two distinct ways. This dimension can manifest where the outcomes of the programme are so accepted, the networks around these outcomes so secure, and the practical delivery of these seen to be so feasible that, even though the particular modality was no longer relevant, the programme outputs could be sustained. For example, the legacy of a conditional cash transfer scheme to encourage girls into school might have resulted in parents and communities valuing girls' education, journeys to school feeling safer, and school and curriculum design becoming more girl-friendly, with the result that the aims of the programme are sustainable even if the modality (CCT) was not. To explore this possibility, the evaluation might conduct a social network analysis, focus group discussions to determine acceptability, a governance review, and a force-field analysis. This dimension and Dimension E 'Social Transformation' are very similar. The main distinction has to do with coverage or scope. Consolidation would include the above change is limited to those who were directly affected by the initial intervention, whereas social transformation presumes a level of transformation that includes individuals and groups who were not directly part of the intervention.

The second possible manifestation of consolidation is where the intervention is recognised as valuable, where the perpetuation of an activity is relevant

(replication), but where its value lies in the expansion (multiplication of the effort). This dimension suggests that the intention behind the intervention does not change but that the implementation of it does. For example, the inclusion of a new curriculum across an education system could be understood as consolidation. The integration of the material would not require any changes in the way the sector operates (see sectoral transformation, dimension D) but rather constitutes the multiplication of using a particular tool, or expansion of an activity or intervention modality.

Consolidation, as related to scale up, requires that the necessary resources and conditions be in place for the scaling up to be sustained. This dimension of sustainability is not the simple scale up of an activity but the assurance that the conditions exist to ensure the scale up can be maintained.

Dimension D Sectoral transformation: the dimension of sustainability materialises when the way the sector works is transformed (e.g. new cross-organisational information flows, collaborations, trust, incentives).

Dimension D, or sectoral transformation, refers to instances where the intervention has led to a redesigned sector re-shaping how and what outcomes are achieved and or how programmes are implemented within a system. For example, a programme designed to take children with parents out of orphanages, return them to their home communities, train social workers to support them and their parents, might result in the transformation of the whole childcare sector to provide new and better ways to meet the needs of children. To demonstrate that the intervention is supporting this type of sustainability, elements that may not have been there before would need to be visible. In the example provided, this would mean the absorption of the intervention model into a systemic social protection system. Another example would include where a curriculum change fundamentally changes how children are taught, and how progress is assessed, with consequences for the whole way education is delivered. This, unlike the example provided in dimension B, and dimension C suggests a retraining of teachers, the development of new testing approach, and new systems to monitor performance of children as well as of schools. This dimension of sustainability is secured deliberately and often requires a wide range of interventions addressing different parts of the system.

Dimension E Social transformation: the aims and structures of social life are adapted changing what is valued, as well as how value is created at a societal level.

Dimension E, or social transformation, is best characterised as a fundamental change in how society views the problem and what should be done to address it. For example, in evaluating efforts to combat Female Genital Mutilation (FGM) we might evaluate whether the programme can be sustained (replication), the aims of reducing incidence delivered into the future (consolidation or sectoral transformation), or the role of midwives, public bodies and mosques transformed

(sectoral transformation). Such a way of examining sustainability would fall under one of the dimensions detailed earlier. Or we might take the view that the long-term sustainability efforts to end FGM depend upon a fundamental transformation in how the rights of women and girls are viewed and achieved through the empowerment of women (social transformation). Evidence for social transformation would be at the level of cultural shifts, legal changes, enforcement activities, and empowerment among women and girls. Another example of this dimension of sustainability would be evidence of a clear focus on reducing inequality by demonstrating a greater priority to inequality through changes to taxation, the labour markets and education systems, and the ability of the population to effectively benefit from these system changes.

Critically the different dimensions of sustainability described earlier show that in each instances sustainability alone is not automatically positive. Rather the achievement of a particular dimension of sustainability must be accompanied by a process of problematisation (i.e. is the dimension of sustainability the 'right' dimension for the intervention under evaluation). Some of the key issues embedded in this problematisation are discussed in the next subsections.

Relevance and sustainability – twin criteria

The dimensions of sustainability we presented are very closely aligned with questions of relevance. Specifically, key questions such as: is what is being perpetuated relevant (and relevant to whom)? Is a critical question that needs to be answered to know if the dimension of sustainability is appropriate to begin with. In other words, should a particular intervention be replicated, consolidated, or should it lead to system transformation, or should it aim for social transformation? Or, indeed, should the approach be abandoned entirely?

Here we suggest that examining sustainability, using tools that are designed based on the approach we propose, is not only about determining what dimension best describes what is being evaluated but also if the dimension of sustainability attained is the appropriate one. Or is there sufficient evidence to determine which dimension of sustainability would be appropriate? Is the intervention being implemented in a way that will achieve the most appropriate level of change and the 'right' dimension of sustainability? For example, if it can be demonstrated that an intervention provides superb results, but the resources cannot even secure its replication, this should be pointed out. Evaluations looking at sustainability should be able to provide insights into the degree to which interventions should be expanded, multiplied and transformed, or the degree to which their termination should be ensured. Regarding the latter, instances where sustainability is not desirable should be clearly identified. Bridging or gateway measures should not be seen as a sustainable solution to a problem. For example, the use of quota systems to ensure that women or minority groups are part of political processes

should not be seen as the attainment of change in gender relations or minority inclusion. Rather these interventions should not automatically pursue their own sustainability. They should present an opportunity for discussion and for the identification of interventions that may lead to social transformation where quotas are no longer needed because the value of inclusive political systems is recognised, and society has changed accordingly.

Sustainability and transformation – twin criteria

All dimensions of sustainability, except for replication, require a degree of transformation. This means that attaining sustainability has meant the transformation, at some level, of what was there before. These transformations may be dynamic, meaning that the changes experienced lead to other changes, which in turn mean more changes. The generation of sustainable results at one dimension may lead to the need for other interventions and for the assessment of these. For example, an intervention could lead to consolidation, which could in turn lead to sectoral transformation, which over time could lead to a system transformation. At each dimension, the transformation could lead to different or new requirements within the system. Critical here is to keep in mind that a change of one intervention type within a system could change the system.

In some instances, interventions can, from the start, envision a transformation model that is founded on attaining different dimensions of sustainability. This we understand as interventions that pursue a sustainability trajectory. The intentional pursuit of a sustainability trajectory could result from either a realisation coming from within the sector that the dimension of sustainability secured is insufficient or that it has evolved to a greater level of coverage/scope. Alternatively, a sustainability trajectory could be driven by an external actor, in development aid this could be a donor or NGO, who recognises that the goals of the interventions are insufficient to address the issue of concern. Another external driver for the pursuit of a sustainability trajectory can be society itself, as is described later. For example, the public becomes aware of a successful intervention which is replicated somewhere and demand that it be expanded (consolidated) or it be integrated into a system in such a way that the system is changed, through the process of demanding change society itself changes.

External drivers of sustainability point to a very important element that should not be overlooked when exploring sustainability: who or what drives sustainability? Understanding different dimensions of sustainability does not mean that one level of sustainability is more important or meaningful but rather acknowledges that activities which start with a ‘study’ or a ‘test and invest’ model can support such a trajectory. Critically the trajectory could in some instances skip certain dimensions. For example, it could start with consolidation or could focus on consolidation of certain interventions and system transformation of another set

of interventions in view of joining both efforts at some stage. Most importantly, as sustainability is not static, neither are the different dimensions or the relationship between dimensions.

Sustainability drivers and constrictors and the sustainability mirage – who and what makes things sustainable

What or who drives sustainability needs to be systematically explored as part of any assessment of sustainability and its achievement. Indeed, the drivers of sustainability might also be critical in the lifespan of sustainability: will a particular dimension of sustainability be sustained forever or for a few months, and critically what will determine its lifespan? What is each dimension of sustainability vulnerable to? For example, if we explore the Afghanistan experience, the freedoms that were enjoyed by Afghan women quickly disappeared with the return to power of the Taliban in August 2021. Those who may have argued that the systems had changed and that women had attained new roles would now need to recognise that this was a very fragile change. Similarly, the change in legislation regarding abortion rights in the USA in 2022 would also indicate a fragility of system change sustainability. Prior to the USA supreme court vote, many would have been forgiven for thinking that abortion rights were recognised at a societal level and that the past decades had signalled social transformation. The above examples illustrate on the one hand the fragility of sustainability even at the broadest encompassing dimensions. It also highlights that sustainability at the system level is elusive and what may appear to be sustainable at one moment in time may not appear so at a later moment – how sustainable is sustainable? Is it sustainable or are we seeing a sustainability mirage?

These are all important questions, and the above examples point to instances where what was thought sustainable proved not to be. These experiences demand that evaluators pay attention to sustainability mirages. Closely tied to what is actually sustainable and what is a mirage of sustainability is the identification of those who can support sustainable change. In the above examples, the change in Afghanistan failed to recognise that it was completely predicated on the Taliban not taking power again. Likewise in the United States, abortion rights were predicated on the views and opinions of those sitting in the supreme court. Anyone studying the issue was well aware of its fragility and the focus that a wide range of parties had placed on revoking the supreme court-established right. Therefore, none should have been a surprise.

Given the aforementioned critical challenge of being able to identify the ‘true’ drivers or constrictors of sustainability, drivers and constrictors may not be specifically related to the intervention. Tests of our approach have shown that there can be important relationships between replication, consolidation, and sectoral transformation on the one side and social transformation on the other (see

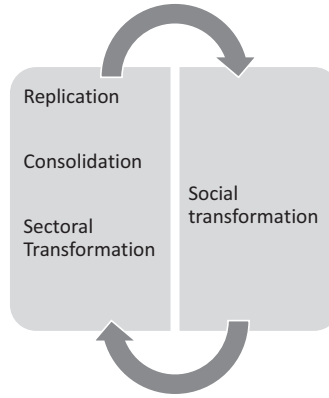

Figure 15.1 Relationship between different dimensions of sustainability.

Figure 15.1). Mainly, that social transformation can drive replication, consolidation, and sectoral transformation or vice versa. For example: campaigns such as the Campaign to Ban Landmines in the 1990s employed the role of society as a driver for change. Likewise, Extinction Rebellion has aimed to generate sectoral transformation. These efforts suggest that individual interventions aiming to address climate change and extinction concerns would not be able to achieve sectoral transformation alone, and that social engagement could be a critical driver. Likewise, it could be argued that a focus on social change could have mitigated the impact that the Taliban regaining of power and the supreme court ruling could have had. Such an effort may have taken longer to see immediate results but may have generated vastly different results.

The relationship between different dimensions of sustainability is not static or linear and is affected by time. Indeed, it may mean that to start with small interventions push an idea, concept, or solution and eventually society takes it over and things take off. Much like rolling a boulder to the mountain top.

Operationalising the different dimensions of sustainability as part of the evaluation process

The approach we introduce here is intended to enable the development of case-specific rubric-based tools which can measure sustainability. Our approach should facilitate the development of tools which enable a more systematic assessment of sustainability and help evaluators navigate different sustainability dimensions. It is intended to provide a framework for how to understand sustainability and enable the development of tools that can effectively identify and

Table 15.1 Matrix identifying the dimensions of sustainability and what is needed to achieve sustainability

Dimension	*What is needed to achieve this level of sustainability?*			
Discovery **Replication** **Consolidation** **Sectoral Transformation** **Social transformation**	**Money:** Financial implications of achieving this dimension of sustainability? Have these been met?	**People:** Personnel implications of achieving this dimension of sustainability? Have these been met?	**Drivers:** Who/what can ensure this dimension of sustainability is reached? Are they/it available?	**Trajectory:** Is this the level of sustainability that should be attained (end game)? Should this dimension of sustainability lead to another?

measure the critical elements of sustainability. It is an aid to thinking rather than a protocol to be followed.

To support this navigation, we have developed a simple tool to facilitate the development of a rubric (and introduced key questions that can facilitate the development of a rubric) to assess the sustainability of the intervention under review. The task of developing the rubric lies with the evaluator and the content should be determined by the thing being evaluated.

Seven questions to consider in developing a rubric-based evaluation

1. From the available data: What dimension of sustainability was pursued?
2. Build a rubric that can respond to the questions in the following matrix and map according to the dimension of sustainability pursued (question 1).
3. Was the dimension of sustainability pursued met?
4. Was a different dimension of sustainability met?
5. Was the most appropriate level of sustainability pursued?
6. Was the most appropriate level of sustainability secured?
7. What should be the lifecycle of the level of sustainability attained? Or was sustainability of what should be sustained attained?

These questions should lead each evaluator to develop a rubric-based tool that allows them to systematically explore sustainability.

These models in practice, and what can be learned from field experience

Examining how sustainability is addressed in a wide range of efforts in the field of development cooperation has shown that most often sustainability aims to secure replication and has limited aspirations beyond that. Evaluations of intervention

models that have aspired to be consolidated and to lead to sectoral level and social level change have shown interesting results. A review of a model focused on the reduction of child labour showed that even though the model was successful, and that where it was implemented, it garnered considerable social support (see Figure 15.1), its ability to lead to sectoral changes was limited at best. This was not a result of shortcomings with the model, or with the way it has been consolidated in the different contexts. Rather the experience highlighted that:

A) some types of change can only really be meaningful at a sectoral level. Securing sustainability at 'lower' levels (replication or consolidation) is valuable, but the need is so great that without sector-wide effort, the sustainability is only partially relevant.
B) sectors often are reluctant to change and, even when willing, sector-level change has considerable implications that can take time and a wide range of commitments (funding, staffing, etc).
C) the relationship between society and other levels of sustainability is highly dependent on the opportunities that society has to make demands. This opens questions of democracy and rights and ability to secure basic goods and services. It challenges the notion that lack of social support for an intervention is based on a disregard for the effort, as it may be equally tied to an understanding that there are limited opportunities to make demands, or that demands made will not be heard (Millard et al., 2015).

Concluding note

This chapter problematises the assessment of sustainability and aims to bring some order to the discussion to enable the practical, real-world, evaluation of sustainability in a meaningful way. The chapter is not a panacea, nor does it intend to be one. Rather it intends to provide an overview of key concepts and complexities that influence sustainability and need to be considered in the assessment of sustainability. It also provides a multidimensional framework that can be used to develop tools to assess sustainability. In so doing, this chapter takes aim at the practicalities of evaluating sustainability and hopes to facilitate evaluation practice.

We consider that the urgent importance of addressing sustainability has led to a failure to adequately operationalise the concept within evaluations. Key to this failure is that many approaches overly homogenise what are in fact very different dimensions that are brigaded, together under the name of 'sustainability. These dimensions vary and involve different kinds of evidence to judge their success. One route out of this difficulty is to apply a multi-dimensional and rubric-based approach outlined here, which can be used to develop a case-specific rubric tool.

Notes

1 Agenda21
2 www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm accessed 26/01/2023
3 Personal research by the authors.

References

Gallie, W.B. (1964). Essentially contested concepts. In Gallie, W.B. (ed.) *Philosophy and the Historical Understanding*. Chatto & Windus, London, pp. 157–191.

Millard, A., K. Forss, A. Basu, B. Kandyomunda, C. McEvoy, and A. Woldeyohannes (2015). *Is the end of child labour in sight? A critical review of a vision and a journey*. Stop Child Labour, Hivos, Kinderpostzegels, Geneva.

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Resolution adopted by the General Assembly on 25 September 2015, General Assembly. A/RES/70/1, Geneva.

World Commission on Environment and Development. (1987). *Our common future*. Oxford University Press, Oxford.