

https://dspace.org/

# Voluntary National Reviews in Africa guide

**CLEAR-AA; UNICEF**

**2019-11**
CLEAR-AA
https://hdl.handle.net/10539/28697

Downloaded from a DSpace repository. https://dspace.org/

clear
Centers for Learning on Evaluation and Results
ANGLOPHONE AFRICA

unicef
for every child

# EMBEDDING EVALUATION IN VOLUNTARY NATIONAL REVIEWS IN AFRICA: A GUIDE



NOVEMBER 2019

clear
Centers for Learning on Evaluation and Results
ANGLOPHONE AFRICA

unicef
for every child

# EMBEDDING EVALUATION IN VOLUNTARY NATIONAL REVIEWS IN AFRICA: A GUIDE

NOVEMBER 2019

**Submitted by:**
Ian Goldman
CLEAR Anglophone Africa
University of Witwatersrand

**Submitted to:**
Ada Ocampo
Senior Evaluation Specialist, Evaluation Office
UNICEF

## Copyright

Copyright of this guideline is vested in UNICEF, while recognizing that many of the concepts come from other sources. In general, publication of excerpts is to be welcomed, subject to acknowledgement of the source.

How to cite this report: UNICEF/CLEAR, Embedding Evaluation in Voluntary National Reviews in Africa: A Guide, UNICEF, New York, November 2019.

## Disclaimer

It should be noted that any opinions expressed in this report are the responsibility of CLEAR Anglophone Africa and not necessarily UNICEF.

## Acknowledgements

This report was independently prepared by Ian Goldman, Candice Morkel and Ian Hopwood of CLEAR Anglophone Africa for the UNICEF Evaluation Office. Ada Ocampo of the UNICEF Evaluation Office provided detailed in-depth inputs, with contributions from Asela Kalugampitiya, George Laryea-Adjei, Michele Tarsilla and UNICEF more broadly.

# CONTENTS


Acronyms and abbreviations 5

PART A: INTRODUCTION 6

1 Background 7

1.1 Who is this guide for? 7

1.2 Background to the voluntary national review process 7

2 What is evaluation and why it is important? 12

2.1 What is evidence? 12

2.2 What is evaluation? 14

2.3 Why is evaluation important? 15

2.4 Overview of types of evaluation 17

3 How can evaluation evidence be embedded in country policies and programmes, the domesticated Sustainable Development Goals and in voluntary national review reporting? 19

3.1 Overview 19

3.2 Using evaluation evidence for diagnosis 19

3.3 Using evaluations to inform planning and design of new policies and programmes 20

3.4 Using evaluations to assess and inform implementation 22

3.5 Evaluating achievement and outcomes 22

PART B: GUIDANCE ON EVALUATIONS AND EVALUATIVE TOOLS 25

4 Different types of evaluation 26

4.1 Types of evaluation 28

4.2 Rapid evaluative methods 33

4.3 Deciding what type of evaluative process to undertake 36

5 Other evaluative tools 37

5.1 Importance and use of a theory of change 37

5.2 Creating an evaluation repository 39

5.3 Develop an evidence (gap) map using evaluations 40

5.4 Synthesizing evidence from existing evaluations 42

PART C: GOING BEYOND EVALUATION TO EVALUATION ECOSYSTEMS 44

6 Evaluation ecosystems 45

6.1 What is an evaluation system or ecosystem? 45

6.2 Practical ways to build the evaluation system using the voluntary national review process 47

PART D: APPLYING THIS IN THE VOLUNTARY NATIONAL REVIEWS 48

7 Applying this in the 2020–2021 voluntary national reviews 49

7.1 Applying this in the 2020 voluntary national review 49

7.2 Planning for the 2021 voluntary national review cycle 51

7.3 Support available 53

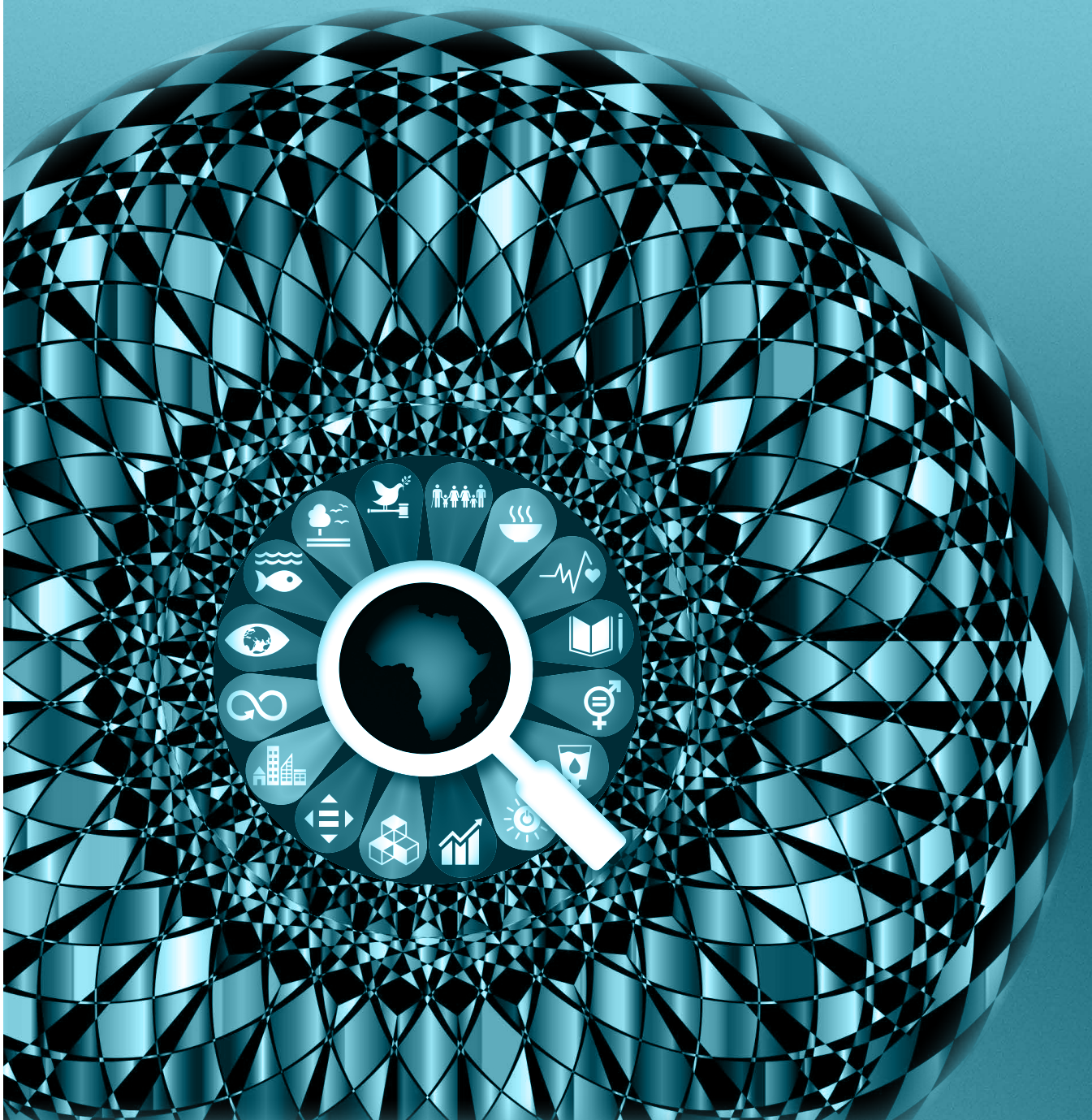
ANNEX 1: REFERENCES 54

## ACRONYMS AND ABBREVIATIONS

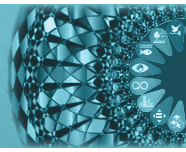
| | |
|---|---|
| **3ie** | International Initiative for Impact Evaluation |
| **AFREA** | African Evaluation Association |
| **APNODE** | African Parliamentary Network on Development Evaluation |
| **CLEAR AA** | Centre for Learning on Evaluation and Results – Anglophone Africa |
| **DPME** | Department of Planning, Monitoring and Evaluation (South Africa) |
| **SDG** | Sustainable Development Goal |
| **UNICEF** | United Nations Children's' Fund |
| **VNR** | Voluntary national review |
| **VOPE** | Voluntary Organizations of Professional Evaluation |
| **WACIE** | West African Capacity Building and Impact Evaluation |

# PART A

## INTRODUCTION



# 1 BACKGROUND



## 1.1 WHO IS THIS GUIDE FOR?

This guide is intended for officers in African governments who are involved in developing voluntary national reviews (VNRs) of country performance against the Sustainable Development Goals (SDGs). It is also intended for agencies supporting VNR processes, such as UNICEF country offices, donors, etc.

It is important that the SDGs are part of government planning, and not seen as external goals separate from what governments and other actors are doing. Hence the VNRs should not be separate processes, but report on what governments and other actors are doing to address their domestic goals and their links to the SDGs. For the VNR process to be valuable, it should not just be a compliance exercise, but contribute to reflection, learning and improvement of government and non-government programmes, and assist in integrating the SDGs into these domestic goals.

For the VNRs to be meaningful, they need to be based on evidence. Evaluations are a powerful source of evidence of how or how not government policies and programmes are working and why, and guidance on how and where to improve. Evaluations demonstrate where resources are being poorly used, and enable performance to be improved using the same budget envelope.

The objective of this Guide is therefore to assist in the incorporation of evidence from evaluations to inform country policies and programmes, and the use of these results in the development of the VNRs.

The Guide includes four parts. Part A, provides an introduction to the VNR process (section 1.2), introduces what evidence and evaluation are and why they are important (section 2) and how evidence can be embedded into policy and practice (section 3). Part B provides guidance on evaluation types and tools, including rapid evaluations (sections 4 and 5). Section 5 includes an introduction to theories of change and how existing evaluation evidence can be used. Part C discusses the evaluation ecosystems that need to be promoted to support evaluation (section 6). Finally, Part D (section 7) illustrates how evaluative processes can be incorporated into the 2020 and 2021 VNRs and provides contact details for follow up.

## 1.2 BACKGROUND TO THE VOLUNTARY NATIONAL REVIEW PROCESS

The 2030 Agenda for Sustainable Development (2030 Agenda) was adopted by world leaders in September 2015 at a historic United Nations Summit.[1] On 1 January 2016, the 17 SDGs officially came into force. The SDGs aim to end all forms of poverty and call for action by all countries – poor, rich and middle-income – to promote prosperity while protecting the planet. They recognize that ending poverty must go hand-in-hand with strategies that build economic growth and address a range of social needs, including education, health, social protection and job opportunities, while tackling climate change and environmental protection.

The Addis Ababa Action Agenda that came out of the Third International Conference on Financing for Development[2] provides concrete policies and actions to support the implementation of the new agenda.

[1] United Nations, 'Transforming our World: The 2030 Agenda for Sustainable Development', Sustainable Development Goals Knowledge Platform, https://sustainabledevelopment.un.org/post2015/transformingourworld.

[2] United Nations Department of Economic and Social Affairs, 'Third International Conference on Financing for Development', www.un.org/esa/ffd/ffd3/index.html

Implementation and success were to rely on countries' own sustainable development policies, plans and programmes, and led by countries. The SDGs were to be a compass for aligning countries' plans with their global commitments.

The follow-up and review processes were to be informed by an annual SDG progress report to be prepared by the Secretary-General. The annual meetings of the High-Level Political Forum on Sustainable Development are supposed to play a central role in reviewing progress towards the SDGs at the global level. The means of implementation of the SDGs are supposed to be monitored and reviewed.[3]

African governments have developed their own long-term development agenda – the 2063 Agenda – which is "a strategic framework for the socio-economic transformation of the continent over the next 50 years. Its builds on, and seeks to accelerate, the implementation of past and existing continental initiatives for growth and sustainable development." The seven aspirations in the 2063 Agenda are:

1. A prosperous Africa, based on inclusive growth and sustainable development;
2. An integrated continent, politically united, based on the ideals of Pan-Africanism and the vision of the African Renaissance;
3. An Africa of good governance, democracy, respect for human rights, justice and the rule of law;
4. A peaceful and secure Africa;
5. An Africa with a strong cultural identity, common heritage, values and ethics;
6. An Africa whose development is people driven, relying on the potential offered by people, especially its women and youth and caring for children; and
7. An Africa as a strong, united, resilient and influential global player and partner.

As well as domesticating the SDGs, countries are expected to domesticate the integration of the First Ten Year Implementation Plan for Agenda 2063 into their national plans.

Governments have committed to implement the SDGs and adapt and incorporate them into their national planning processes. As stated in paragraph 84 of the 2030 Agenda, regular reviews in the High-Level Political Forum on Sustainable Development[4] are to be voluntary, country-led, undertaken by both developed and developing countries, and provide a platform for partnerships, including through the participation of major groups and other relevant stakeholders. VNRs make possible the sharing of experiences, including successes, challenges and lessons learned, with a view to accelerating the implementation of the 2030 Agenda.

The process of carrying out the VNR should not be seen as separate from the implementation of the SDGs. Rather than an end in itself, the VNR is a process by which countries take stock of and assess progress and shortcomings in the implementation of the SDGs and their targets.

In order to take stock, evidence is critical to find out what is working and what is not, for whom and under what circumstances. Evidence will also help identity what needs to be done to accelerate the attainment of the different SDGs. In this vein, evidence generated through evaluations is key.



[3] Definition of Agenda 2030 drawn from United Nations, 'The Sustainable Development Agenda', Sustainable Development Goals, www.un.org/sustainabledevelopment/development-agenda/ .

[4] United Nations, 'High-Level Political Forum on Sustainable Development' Sustainable Development Goals Knowledge Platform, https://sustainabledevelopment.un.org/hlpf .

The VNRs are intended to track progress towards implementing the SDGs and targets in a manner that respects not only their universal and integrated nature but also all dimensions of sustainable development. The principles guiding follow-up and review at all levels, contained in paragraph 74 of the 'Handbook for the Preparation of Voluntary National Reviews', provide, among other things, that reviews will be substantive and knowledge-based, as well as open.[5] The 2030 Agenda specifies that the purpose of follow-up and review is to ensure accountability and that follow-up and review mechanisms should be inclusive, participatory, transparent, people-centred, gender-sensitive, respect human rights and have a particular focus on the poorest, most vulnerable and those furthest behind.

If the SDGs/VNRs and Agenda 2063 are to contribute to development in Africa, the targets that are domesticated must be meaningful, encourage performance in the areas where countries are focusing, and encourage learning from the lessons of implementation.

Key to this is that the targets and plans and reporting against these are evidence-based. If governments draw on the best available and most relevant evidence, they are more likely to make successful decisions. That does not mean they will always accept that evidence, but they will at least have considered it. That is the basis of evidence-based policy-making and implementation or evidence-informed decision-making. Figure 1 provides a theory of change for how using evidence from evaluations and other sources can improve development outcomes. An orientation course was held in Addis Ababa in December 2018 that provided examples of how this can work using research, data and evaluations.

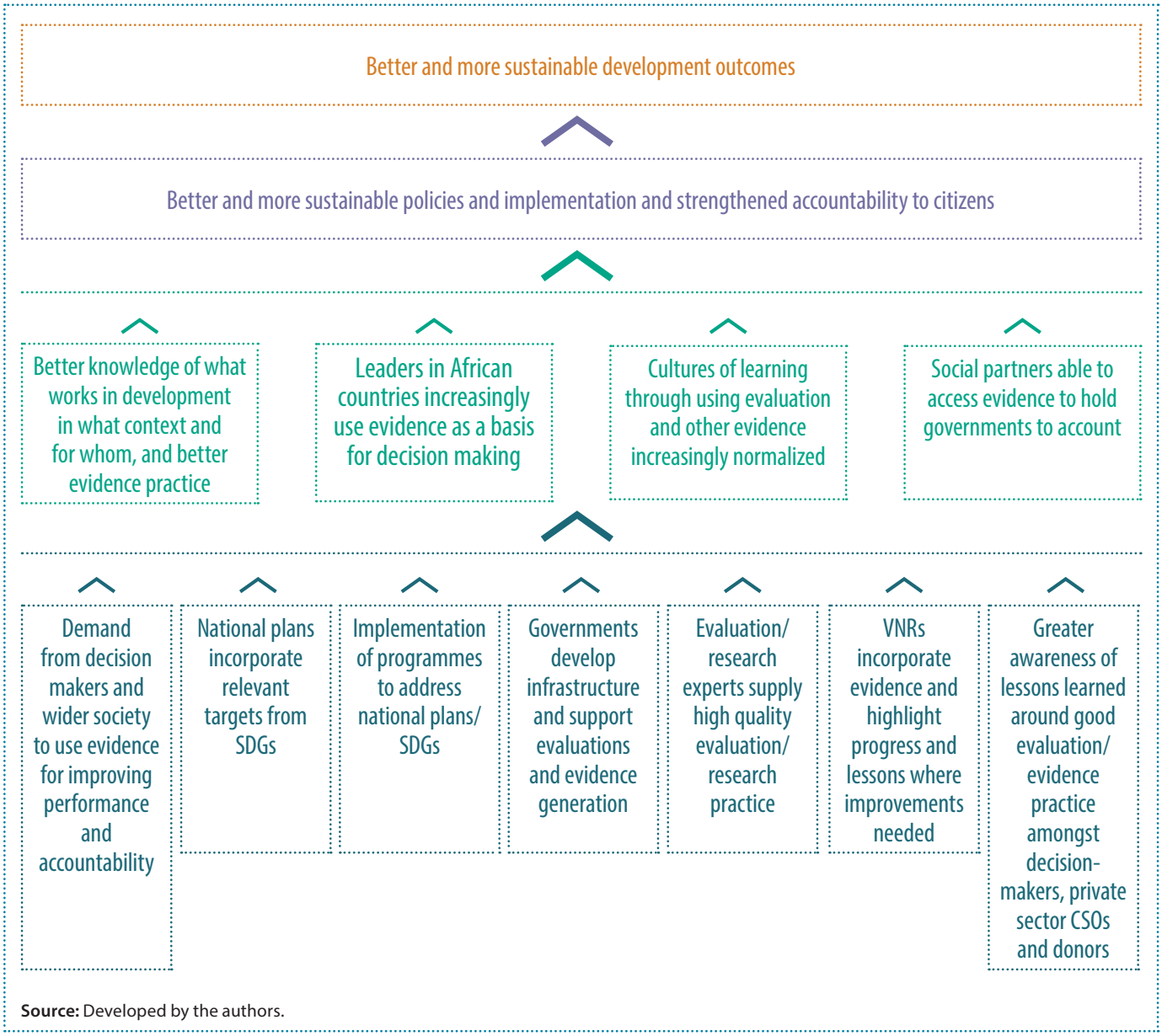
A narrative version of the theory of change is based on a common endeavour to improve development outcomes. To achieve this, there is a need for better and more sustainable policies and implementation and strengthened accountability to citizens. This requires evidence; that is, better knowledge of what works and for whom and leaders in Africa increasingly using this evidence, with cultures of learning through using evidence becoming normalized. To ensure accountability and facilitate the involvement of a wider range of stakeholders, social partners must be able to access evidence and hold government accountable. For this to happen, there needs to be demand for evidence, incorporation of SDG targets into national plans, the establishment of an evidence generation infrastructure, the supply of high-quality evidence, and the incorporation of that evidence into the VNRs. In the process, it is important to build greater awareness of lessons around good evaluation practice in the country.

The guidance for countries on how to implement their VNRs comes from the handbook issued each year. This builds on the Secretary-General's voluntary common reporting guideline for VNRs, which is annexed in each annual handbook.[6]

[5] United Nations Department of Economic and Social Affairs, 'Handbook for the Preparation of Voluntary National Reviews', DESA, 2018, https://sustainabledevelopment.un.org/content/documents/20872VNR_hanbook_2019_Edition_v2.pdf

[6] Ibid.

## Figure 1: Theory of change for use of evidence leading to better development outcomes




**Source:** Developed by the authors.

## Secretary-General's guidelines for voluntary national reviews:

Key elements are:

a. They will maintain a longer-term orientation, identify achievements, challenges, gaps and critical success factors and support countries in making informed policy choices. They will help to mobilize the necessary means of implementation and partnerships, support the identification of solutions and best practices and promote the coordination and effectiveness of the international development system.

b. They will be rigorous and based on evidence, informed by country-led evaluations and data which is high-quality, accessible, timely, reliable and disaggregated by income, sex, age, race, ethnicity, migration status, disability and geographic location and other characteristics relevant in national contexts.

c. They will require enhanced capacity-building support for developing countries, including the strengthening of national data systems and evaluation programmes, particularly in African countries, least developed countries, small island developing States, landlocked developing countries and middle-income countries."[7]

The 2019 edition of 'Handbook for the Preparation of Voluntary National Reviews' notes:

> "Draw on existing reports. Use existing national platforms and processes that could contribute to the VNR writing and analysis process. Examples include:
>
> - National frameworks such as national development plans and national sustainable development strategies;
> - Reports submitted to international bodies, including human rights mechanisms, such as the Universal Periodic Review (UPR) and international treaties, and other reports such as the Nationally Determined Contributions (NDCs) relating to the Paris Agreement;
> - The SDG reports prepared at the national level together with the United Nations Country Team (UNCT) and the United Nations Development Programme (UNDP) if available or applicable."[8]

These existing reports should also include existing research and evaluations.

This Guide introduces how evidence can be used in different stages of the programme/policy cycle. It does not cover the main planning tools, except one – the theory of change – which is a key tool used in evaluation that should be carried out as the first step in a planning process.

This Guide gives an overall approach to using evidence, concentrating on how to use evaluation evidence and related tools to support VNRs.

[7] Ibid, pp. 57–58.

[8] Ibid, p. 15.

# 2 WHAT IS EVALUATION AND WHY IT IS IMPORTANT?

## 2.1 WHAT IS EVIDENCE?

Evaluation is one source of evidence to inform policy and practice. Here evidence is considered first. According to Cairney (2016), "evidence is assertion backed by information."[9] The Programme to Support Pro-Poor Development in South Africa (2011) defines evidence as data that meet the standards of:

- **Internal validity** – the extent to which the design and conduct of the study eliminates the possibility of bias (is it true internally of the case in a particular study);
- **Adequacy of reporting** – are the statistics adequate and does data support the findings?
- **And potentially external validity** – can the study be extended (is it relevant) to the wider world (e.g., can a pilot be scaled up?)[10]

In this definition, the key qualities of evidence are independence, objectivity and being verifiable. In contrast, opinions are statements and claims that do not meet the standards of evidence and are positional, subjective, partial (selective) and hard to verify.[11]

Evidence can come from many sources, including the following:[12]

- **Statistical evidence** from surveys, official statistics and administrative data, each of which can indicate the size, nature and dynamics of the problem at hand;
- **Descriptive and experiential evidence,** which illuminate the nature, size and dynamics of a problem;
- **Individual evaluations and research studies;**
- **Research synthesis,** including systematic reviews of evidence, meta-analyses and rapid evidence assessments;
- **Economic and econometric evidence,** which refers to the cost benefit and effectiveness of interventions;
- **Implementation evidence** indicating how similar policies have been successfully implemented, and how barriers to successful implementation have been overcome;
- **Ethical evidence** in terms of questioning or understanding the ethical implications of the policy, including issues of equity, fairness and justice.

Figure 2 gives a picture of a policy/programme cycle developed for training in evidence-based policy and implementation in Africa that includes the stages of agenda setting, diagnosis, selection of intervention, planning/design, implementation, evaluation and ongoing learning.13 Evidence can be used around different stages of the policy cycle for diagnosis (e.g., the size of a problem indicated by a community survey), implementation (progress) or the outcomes (e.g., the effectiveness of a solution). Evaluative processes can be used at different stages of this.

[9] Cairney, P., 'The Politics of Evidence-Based Policy Making', Palgrave Macmillan, London, 2016, p. 4.

[10] Programme to Support Pro-Poor Development in South Africa, 'Evidence-Based Policy-Making (EBPM): Enhancing the use of evidence and knowledge in policy making', PSPPD, July 2011.

[11] Training materials used in training of senior managers in evidence at the University of Cape Town, 2013–2018.

[12] Adapted from 'Evidence-Based Policy-Making (EBPM)', p. 2.

[13] This cycle is derived from training materials on evidence used in a range of settings, including the University of Cape Town/Department of Planning, Monitoring and Evaluation course for senior managers.

## Figure 2: Policy/programme cycle




**Source:** University of Cape Town training materials used for senior managers.

For further information on evidence-based policy-making, see 'Evidence-Based Policy-Making (EBPM)', a policy brief by the Programme to Support Pro-Poor Policy Development in South Africa.[14]

[14] Available at http://psppdknowledgerepository.org/component/jdownloads/download/7-evidence-based-policy-making-ebpm-enhancing-the-use-of-evidence-and-knowledge-in-policy-making (accessed 21 August 2019).

## 2.2 WHAT IS EVALUATION?

**Monitoring** tracks whether you are doing what you planned at the output and outcome levels, while **evaluation** seeks to see why intended and unintended outcomes are being achieved, the quality (merit, worth and significance) of the achievements of a policy, programme, etc. and how you can improve. Box 1 provides the United Nations Evaluation Group definition of evaluation.

The **purpose** of evaluations can be the following:

- **Accountability:** For example, where is public spending going and what public value is there from investment? Accountability can be to Parliament, to broader society or to senior managers. For example, is this spending making a difference? However, this can quickly become a compliance exercise unless it is also used for learning.
- **Learning:** Evaluation is used to inform continuous improvement and the key clients are the managers themselves, to understand how to strengthen their work, potentially as clients or users of services.
- **Improving decision-making:** Questions could be, for example, should the intervention be continued? How should implementation be changed? Should budget be increased/decreased?
- **Improving knowledge:** By systematically storing evaluations in repositories, this information can be accessed by those seeking to design new policies or programmes, or in the academic space to understand the work of sectors.
- **Improving governance:** Monitoring and evaluation and accountability are key aspects of good governance. Effective governance requires knowledge of what is working or not, and feedback for improvement.

In practice, evaluation may serve all of the above-mentioned purposes; and all of these purposes can be taken into account when it comes to evaluation systems. Table 1 shows how evaluation can be differentiated from other evidence processes.

### Box 1: United Nations Evaluation Group definition of evaluation

An evaluation is an **assessment**, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of **achievement** of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using appropriate **criteria** such as relevance, effectiveness, efficiency, impact and sustainability.

An evaluation should provide **credible**, useful evidence-based information that enables the timely incorporation of its findings, **recommendations** and lessons into the decision-making processes of organizations and stakeholders.

**Source:** United Nations Evaluation Group, 'Norms and Standards for Evaluation', UNEG, New York, 2016.

## Table 1: Comparison of evaluation and other evidence tools

| | ACTIVITY | OBJECTIVE |
|---|---|---|
| **Inspection/ investigation** | An inspection/investigation detects wrongdoing and verifies information. | Control and compliance |
| **Audit** | An audit is the assessment of a person, organization, project or product primarily for the purpose of determining its validity and authenticity, or to verify adherence to a set of pre-defined processes. | Compliance with procedures |
| **Performance audit** | This is an independent auditing process to assess measures instituted by management to ensure that allocated resources are procured economically and utilized efficiently and effectively and, if necessary, to report thereon (Auditor General of South Africa). | Accountability, control and compliance |
| **Monitoring** | Monitoring is ongoing tracking of progress against plans and initiating corrective action. | Management, accountability and corrective action |
| **Evaluation** | Evaluation is systematic assessment of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance to analyse the level of achievement of both expected and unexpected results and causality using appropriate criteria such as relevance, effectiveness, efficiency, impact and sustainability to provide recommendations and lessons for decision-making. | Understand/learn, accountability, improving performance, inform policy, planning and budgeting |
| **Research** | Research involves testing hypotheses/propositions through observation of reality. | Learning/knowledge creation only (no accountability focus), can inform policy |

**Source:** Adapted from South Africa's National Evaluation Policy Framework, 2011.

## 2.3 WHY IS EVALUATION IMPORTANT?

Without evaluation, there is no rigorous reflection on whether the right things are done in the right ways, what is or is not working and why so. Without this, there is a risk of repeating mistakes, wasting scarce funds and loosing impact on citizens. In practice, work on evaluation in Africa has shown that many programmes and policies are poorly planned, often with inadequate diagnosis of the root causes and how to address these. The theory of change of how the intervention will actually be implemented is often poorly thought through, and it takes the evaluation to uncover this and help with re-planning the intervention. In the absence of evaluation, mistakes will continue to be made and new projects and programmes will be designed without learning from the past.

In practice, the culture of many African governments is not very conducive to learning. While Benin, South Africa and Uganda are considered among the foremost exponents of evaluation in Africa, a recent study of the performance culture in these three countries found that about 50 per cent of managers working there do not see problems as an opportunity for learning. The study found that about 30 per cent of managers see resistance to the use of evidence as a problem. The situation is probably more acute in countries where evaluation is not well established.

## Table 2: The major barriers to the effective use of evaluation in decision-making, learning and accountability in your department

| BARRIER | % OF RESPONDENTS SAYING ALWAYS/OFTEN (BY COUNTRY) | | |
|---|---|---|---|
| | BENIN | SOUTH AFRICA | UGANDA |
| The concealing of findings is a barrier to the effective use of monitoring and evaluation | 24.2 | 31.7 | 34.7 |
| Resistance from senior management to transparent decision-making processes | 35.6 | 27.9 | 33.3 |
| Little respect for evidence-based decision-making in the department | 30.9 | 27.9 | 34.7 |
| The hierarchy makes it difficult to openly and robustly discuss performance | 40.3 | 38.5 | 42.7 |
| Managers fear admitting mistakes or problems | 49.0 | 54.8 | 46.7 |
| The monitoring and evaluation unit has little influence in the department | 45.6 | 51.9 | 48.0 |
| Monitoring and evaluation is seen by management as policing and a way of controlling staff | 43.0 | 44.2 | 37.3 |
| Monitoring and evaluation is regarded as the job of the monitoring and evaluation unit, not of all managers | 63.8 | 54.8 | 54.7 |
| Senior management does not champion monitoring and evaluation and there is no honesty about performance | 40.3 | 41.4 | 34.7 |
| Problems not treated as an opportunity for learning and improvement | 45.0 | 40.4 | 46.7 |
| There is no consistent demand for evaluation from ministers and management | 28.2 | 23.1 | 32.0 |
| Concerns exist with respect to unhelpful conclusions on policies' effectiveness | 42.3 | 52.9 | 50.7 |

**Source:** Goldman, et al., 'Mere compliance or learning – M&E culture in the public service of Benin, Uganda and South Africa', chapter 4 in Goldman, I and Pabari, M (eds) 'Policy Making in Africa: Reflections on Using Evidence', Routledge, forthcoming 2020.

All of the barriers described in Table 2 illustrate challenges in the use of the VNR process in a reflective way to help inform improvements in country performance. In later sections, the Guide suggests ways to address this. Evaluation is a key tool for understanding whether the outcomes planned in country targets and the SDGs are being achieved and why, and how interventions need to be strengthened to maximize the likelihood of attaining the targets. Without evaluations, one may know whether there is improvement from outcome monitoring (e.g., is life expectancy increasing), but not why that improvement is happening or how to make it happen faster, better or for specific groups.

## 2.4 OVERVIEW OF TYPES OF EVALUATION

Evaluations can be described in several ways. Key ways of looking at evaluations include:

- **When** the evaluation happens: before, during or after an intervention, leading to terms such as ex-ante, process and ex-post evaluation;
- Whether the main **purpose** is to learn (formative) or to judge performance (summative);
- By the key **question**: how is the intervention working (implementation evaluation), is the intervention having an impact (impact evaluation), what is the cost-benefit (economic evaluation);
- The main **methodology** used (qualitative, quantitative, experimental, quasi-experimental, mixed method);
- By approaches e.g., realist (key focus on context and how certain mechanisms result in outcomes), theory-based (test the theory of change), developmental (accompany the intervention) and empowerment (key focus on empowerment of stakeholders), etc;
- By the **time available**, which defines the scale and degree of formality and rigour needed.

Based on this overview, a set of evaluation types can be described:

- **Ex-ante/prospective evaluation/diagnostic review:** This supports the design of interventions. Its purpose is to gather information and carry out analyses that help to define objectives, to ensure that these objectives can be met, that the options used are cost-effective and that reliable later evaluation will be possible.
- **Process/implementation/mid-term evaluation:** This evaluates whether an intervention's operational mechanisms do or do not support the achievement of the objectives and why. It looks at activities, outputs and outcomes, the use of resources and causal links. It is used to improve the efficiency and efficacy of operational processes. It can be rapid, primarily using secondary data, or in-depth with extensive field work. This may be applied once or several times during the operation of the intervention.

- **Ex-post (or retrospective) evaluation:** This assesses/evaluates quality towards the end of a programme or after a programme or institution has been in operation to determine strengths and weaknesses. This could look essentially at implementation, at outcomes, at impact or at all of these.
  - **Outcome evaluation:** This assesses achievement of the intended or unintended outcomes that a specific intervention contributes to.
  - **Impact evaluation:** This measures changes in the well-being of the target population attributable to a specific intervention, or to which an intervention contributes.[15]
  - Both outcome and impact evaluations are used to inform high-level officials on the extent to which the intervention should be continued or not, and if any potential modifications are needed. Ideally, they should be designed early on, with a baseline implemented early and outcomes checked at three to five years and impact checked at five years or more.
- **Economic evaluation:** This looks at cost effectiveness or cost-benefit and is a key type of evaluation that is often overlooked.

In practice, these can also be combined. For example, a cost-benefit can be combined with an ex-ante to diagnose a situation, what options can be considered for interventions and the cost-benefits of these. This is discussed in more detail in section 4.[16]

In addition to rigorous evaluations, shorter and less rigorous processes can be used where answers are needed rapidly, or where it is not necessary to spend significant amounts of time and budget on the evaluative process. Examples include:

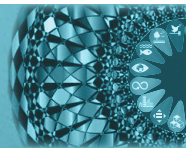
- Two-to-three-day **evaluative workshops**, where carefully selected stakeholders are brought together in a rigorously constructed process to evaluate an intervention collectively;
- **Annual review models**, where a small team reviews an intervention's progress against its theory of change and suggests improvements;
- **Rapid evaluations**, where departmental teams of monitoring and evaluation and content specialists or outsourced service providers do a rapid evaluative process, testing monitoring data and using interviews to understand why or why not things are working and suggest improvements. This approach may also include a workshop.

At a minimum, some reflective processes to understand why could be introduced into monitoring processes (evaluative thinking) so they contribute to adaptive management. Additional details on these processes are provided in section 5.

[15] Note that the term 'impact assessment' is often used to refer to efforts to estimate likely impacts prior to an intervention being carried out (e.g., a regulatory impact assessment or a feasibility study for infrastructure).

[16] For additional resources on evaluation types, visit BetterEvaluation at <https://betterevaluation.org/en/start_here/decide_which_method> or see the guidelines at the South African Department of Planning, Monitoring and Evaluation (DPME) website at <https://evaluations.dpme.gov.za/pages/guidelines-other-resources> (accessed 13 November 2019).

# 3. EMBEDDING EVALUATION EVIDENCE IN COUNTRY POLICIES/ PROGRAMMES, THE DOMESTICATED SDGS AND VNR REPORTING



## 3.1 OVERVIEW

Evidence is needed at different stages of policy/ programme planning and implementation. In the orientation course, this is related to the policy/ programme cycle highlighting some of the types of evidence that can be used at each stage (*see Figure 1*).

1. **Diagnosing**: In understanding the problem, exploring the reasons for poor performance, more livelihood outcomes, etc., in comparing options to respond. This applies to national programmes, as well as to the SDGs.
2. **Planning**: In planning and designing appropriate responses to the diagnosis undertaken, setting indicators and targets, including domesticating the SDGs to fit national agendas, selecting appropriate targets to include and determining what targets to set.
3. **Implementation**: In tracking implementation, using monitoring data, implementation evaluations and implementation research to understand how implementation is going and how it can be strengthened.
4. **Assessing outcomes and impacts:** Outcome monitoring says whether outcome indicators are being achieved (e.g., maternal mortality, passing of school examinations) as well as unintended outcomes, while evaluations seek to understand what caused those outcomes/impacts.
5. **Learning**: Supporting learning throughout the process.

Countries need to consider evaluation in these different phases of their national programming, and this should become part of the normal process of programme planning and management. The VNR reports on the progress made by countries towards implementing the 2030 Agenda and assesses that progress against the 17 SDGs. This is expected to ensure that the SDGs are integrated into national plans, programming and reporting.

## 3.2 USING EVALUATION EVIDENCE FOR DIAGNOSIS

Effective diagnosis is essential to designing suitable policies and programmes. Work undertaken by CLEAR Anglophone Africa (CLEAR-AA), the Twende Mbele Programme[17] and South Africa's DPME all suggest that a key reason for inadequate government performance is insufficient diagnosis, often addressing short-term symptoms rather than underlying root causes of problems. This comes up in training of top-level managers on evidence-based policy-making and the implementation that these partners have undertaken in Benin, South Africa and Uganda and the Orientation Course conducted for UNICEF.[18]

This inadequacy of diagnosis was evidenced in Namibia's 2018 VNR (p. 16), which reported the following for SDG 2: "Weak research and evidence collation combined with programming challenges makes it difficult to quantify the interactions among food and nutrition insecurity, poverty and HIV. There is no food and nutrition security baseline from which to measure progress in addressing hunger. The limited information on food consumption patterns, both overall and by population group, and fragmented

[17] Twende Mbele is an African monitoring and evaluation partnership between governments and regional support agencies promoting the use of monitoring and evaluation to improve development outcomes. The partner countries are Benin, Ghana, South Africa and Uganda, plus CLEAR-AA and the African Development Bank.

[18] For further details on these courses, contact Hannah Diaz at the University of Cape Town (hannah.diaz@uct.ac.za), Cara Waller of Twende Mbele (cara.waller@wits.ac.za) or Ian Goldman (ian.goldman@wits.ac.za).

data collection tools, approaches and systems among sectors and ministries, makes it difficult to obtain a comprehensive outlook of the food and nutrition security situation so as to inform programme design, policy and provide timely government response."

Data is helpful in diagnosis. In explaining root causes, one needs to ask the question why, which evaluation and research are best placed to undertake. Several approaches can be used:

1. **Identifying existing evaluations and research and using their results:** South Africa did an audit of all evaluations in 2011 to build a repository, which currently has over 140 evaluations.[19] Stellenbosch University did a survey in 2016 of evaluations in Africa for CLEAR-AA, which now has a database of over 2,650 African evaluations.[20] The International Initiative for Impact Evaluation (3ie)[21] has a repository of impact evaluations. This is discussed further in section 5.

2. **Synthesising results from existing evaluations:** Once evaluations have been identified, there are methodologies for synthesizing them. These can range from a synthesis of evaluations in a sector (e.g., seven evaluations have been synthesised in the human settlements sector in South Africa), to a systematic search using databases, with strict protocols on relevance and quality for which studies can be synthesised. These produce a range of products with different degrees of synthesis and rigour, from evidence gap maps,[22] to scoping reviews,[23] rapid evidence assessments,[24] systematic reviews, etc.[25] This is discussed further in section 5.

3. **Undertaking new evaluation or research:** If evaluations and research do not exist, new studies may be needed. Research typically falls within the diagnosis phase, as well as a particular type of evaluation – diagnostic evaluations – which seek to understand the current situation and analyse root causes and options. For example, the Diagnostic Review of the State Response to Violence against Women (VAW) and Violence against Children (VAC) in South Africa reviewed both the institutional and programmatic mechanisms by which the state addresses violence against women and children. It considers the whole of government response, covering overarching challenges faced by 11 key departments with roles to address violence against women and children. It considers the state response across the three spheres of government: national, provincial and local.[26] This is discussed further in section 4.

Undertaking evaluation and research requires data. The more that data are disaggregated by age, gender, geography, type of school, etc., the more in-depth analysis can be done, and the more equity-based decisions can be made. Section 3.5 discusses a rapid impact evaluation in South Africa that was possible because outcome data already existed, as did good data on pre-primary attendance.

## 3.3 USING EVALUATIONS TO INFORM PLANNING AND DESIGN OF NEW POLICIES AND PROGRAMMES

In responding to the SDGs, existing policies and programmes may need to be tweaked or new ones developed. So how can countries use evidence to help plan and design policies and programmes?

[19] See <https://evaluations.dpme.gov.za/evaluations.aspx>, accessed 13 November 2019. Benin has a similar repository at <www.presidence.bj/evaluation-politiques-publiques>, accessed 13 November 2019.

[20] See https://crest2.sun.ac.za/african_evaluation_db/default/african_eval_db_01, accessed 13 November 2019.

[21] See www.3ieimpact.org, accessed 13 November 2019.

[22] See www.3ieimpact.org/en/evidence/gap-maps/, accessed 13 November 2019.

[23] See https://guides.library.utoronto.ca/c.php?g=588615&p=4310109, accessed 13 November 2019.

[24] See, for example, <www.cebma.org/faq/what-is-an-rea/, accessed 13 November 2019.

[25] See www.3ieimpact.org/evidence-hub/systematic-review-repository, accessed 13 November 2019.

[26] See https://evaluations.dpme.gov.za/evaluations/2014-02, accessed 13 November 2019.

Some of the VNRs are not explicit on the strategies for achieving the SDGs. Namibia's VNR is explicit (*see Figure 3*).

If one is clear about the strategies to achieve the SDGs, then evaluation is a key tool to see if such strategies are working.

The **theory of change** (sometimes referred to as a logic model or programme theory) is a key tool used in evaluation that should be carried out as the first step in a planning process. If the desire is to achieve a certain number of outcomes, the theory of change will clarify what changes will be needed to get there, with what outputs and activities and based on what assumptions. Having a solid theory of change informed by evidence provides a good start, at least in terms of being clear on how desired outcomes will be achieved. Countries that wish to strengthen their policies and programmes would be well advised to develop theories of change for them, which will also make monitoring and evaluating easier. The stronger the evidence base, the stronger the theory of change. Developing a theory of change requires good facilitation but can be done in one or two days. The countries with established evaluation systems are now including theory of change/programme theory as a key step (*see Table 3*). Section 6 provides guidance on creating a theory of change.

Another key planning step is the setting of **indicators** and **targets**. Evidence is needed to know which indicators and targets are suitable, including the baselines that should be used to measure progress. A key step in relation to Agenda 2030 is the domestication of the SDG targets into national targets. The more that these can be based on evidence, the more realistic they might be. In addition, for some targets, it may be necessary to disaggregate by gender, age, etc., to know that no one is being left behind.

## Figure 3: Extract from Namibia's 2018 VNR (p. 25)

### Key Interventions Supporting SDG 5

**Strategies and targets amongst others:**

| Strategies for the attainment of the targets | Set national targets |
|---|---|
| 1 Increase financial and human capacity of service providers in police, justice, security and education for integrated prevention, protection and response services for victims of GBV, human trafficking and violence. | Reduce GBV prevalence rate from 33 percent (2013) to 20 percent by 2022 |
| 2 Strengthen implementation of gender responsive budgeting and planning to mainstream gender in all sector policies, programmes and budgets of OMAs. | Improve percentage of women in politics and decision making from 47 percent (2015) to 50 percent by 2022 |
| 3 Mainstream informal businesses led by women by ensuring financial inclusion for women in MSME, especially in agri-business, blue economy and extractive industries. | Increase employment rate of women from 68 percent (2014) to 76 percent by 2022 |

### Table 3: Some characteristics of evaluations in Benin, South Africa and Uganda

| COMPONENTS | BENIN | SOUTH AFRICA | UGANDA |
|---|---|---|---|
| Programmes/policies being evaluated have a programme theory | They have a log frame. Theory of change not common. Effort made now for retrospective development of theory of change | Being introduced through the evaluation system. All evaluations now develop theories of change | Some policies and programmes have a log frame. Some newer programmes have theories of change. In all cases, evaluation teams develop the theories of change |
| Type of evaluations | 14 implementation and/or process evaluations, 1 impact evaluation | 45 implementation (process, some summative) evaluations, 8 impact, 5 diagnostic, 1 economic (DPME, 2017a) | 1 implementation evaluation, 4 process evaluations and 3 impact evaluations |

**Source:** Goldman, et al (2018). 'The Emergence of Government Evaluation Systems in Africa'.

## 3.4 USING EVALUATIONS TO ASSESS AND INFORM IMPLEMENTATION

Implementing policies and programmes is the bread and butter work of government. Governments need to understand how their programmes are working and how they can be strengthened to promote wider societal gains. Implementation evaluations seek to understand how interventions are operating, whether the theory of change is working as planned, and whether they are likely to reach the desired outcomes. Often, it is found that implementation is not proceeding as planned, in which case the desired outcomes may not be achieved. The emerging results from implementation/process evaluations can be fed back into practice quickly. Typically, implementation evaluations may be undertaken after two or three years. It is also good to do an annual reflective exercise, for example, an evaluative workshop or annual review to quickly diagnose any emerging problems or changes needed for improvements. Section 5 has some guidance on implementation evaluations.

A key question for evaluation is whether suitable data is being collected to track performance against targets at various levels of the theory of change, for whom and where, which implies that the data needs to be disaggregated. A key partner for this will be national statistical agencies, which should also be able to assist with ensuring the quality of ministerial administrative data.

Countries such as Benin, South Africa and Uganda are predominantly undertaking implementation evaluations (*see Table 3*). These typically take about 12 months from inception to final report approval.

The VNR is essentially reporting on implementation, and depends on monitoring against the targets set. How the SDG and Agenda 2063 targets have been domesticated is therefore important for the VNRs, especially since the International Institute for Environment and Development/EVALSDGs briefs[27] suggest that most countries are using existing monitoring and evaluation systems to report.

## 3.5 EVALUATING ACHIEVEMENT AND OUTCOMES

Traditionally, evaluation has been seen as ex-post, reporting on past performance and explaining the reasons for that performance, often using the Development Assistance Committee criteria of relevance, effectiveness, efficiency, impact and sustainability. This is still an important role

[27] International Institute for Environment and Development, 'Effective Evaluation for the Sustainable Development Goals', IIED, 2016, www.iied.org/effective-evaluation-for-sustainable-development-goals, accessed 13 November 2019.

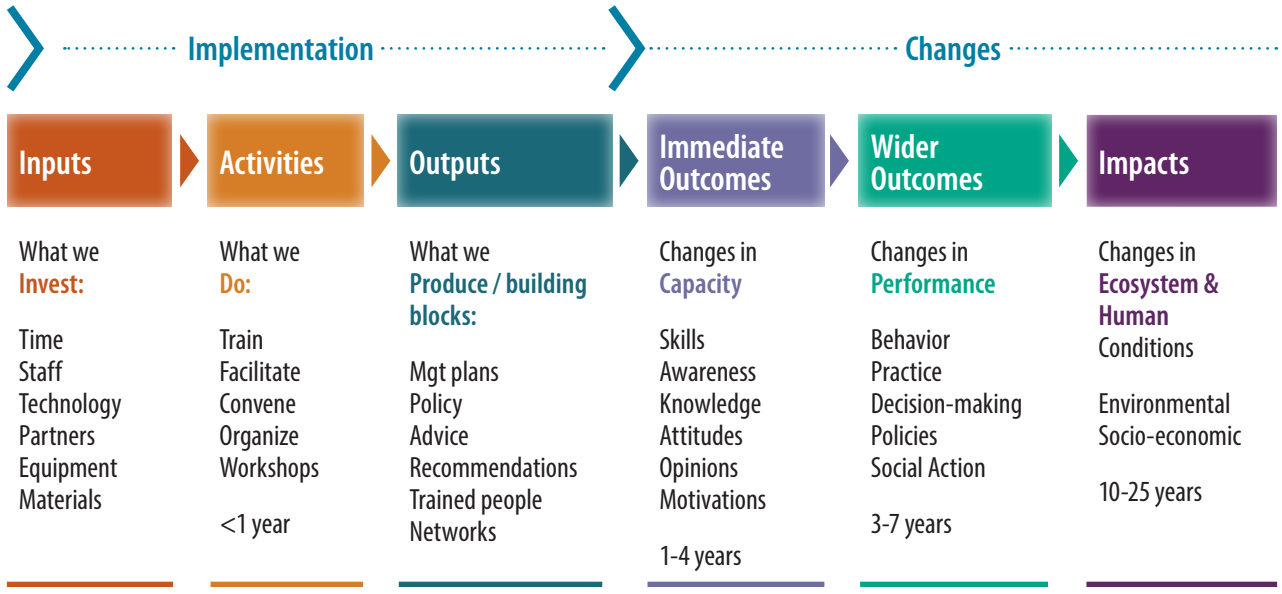
for evaluation, particularly for impact evaluation. However, diagnostic (ex-ante) and implementation/process evaluations are just as important.

Outcome evaluations look at the intermediate changes to impacts. Typically, immediate outcomes would be changes in systems and capacities, while wider outcomes are changes in behaviour and performance (see Figure 4). It is generally easier and quicker to assess these outcomes than the final impacts. Impact evaluations may no longer be necessary if there is a solid outcome evaluation and if there is already good research evidence that proves that if certain outcomes are achieved, specific impacts will result. It is also very important when looking at outcomes to pick up any unintended outcomes, and how the programme logic is playing out to result in these.

Impact evaluation is a powerful tool, and experimental/quasi-experimental impact evaluations can be very authoritative on the effect of the size of a particular treatment or intervention. However, impact evaluations must have a qualitative component if they are to explain why certain impacts are happening or not, or why unforeseen impacts are occurring. Given that impact evaluations should only be undertaken if it is known that implementation is actually happening more or less as planned, undertaking an implementation evaluation first may be appropriate. This can also assess the state of the data. *See Box 2* for an example of the use of results from an impact evaluation in Malawi.

Once again, suitable data is critical. South Africa was able to undertake a very cheap impact evaluation of an additional year of pre-primary education (Grade R) by using data on learners' learning outcomes at over 18,000 schools, and linking that data with data on children that had attended Grade R. The evaluation was very cheap (US$ 20,000) and was carried out very quickly (three to four months) because it was able to use existing data.[28]

## Figure 4: Hierarchy of changes in the programme logic




**Source:** Model developed from the Strengthening Management for Impact Programme led by Wageningen University.

[28] Republic of South Africa Department of Planning, Monitoring and Evaluation, 'Evaluations', https://evaluations.dpme.gov.za/evaluations.aspx?keyword=grade%20R, accessed 13 November 2019.

### Box 2: An impact evaluation of Malawi's cash transfer programme

A 2006 impact evaluation of Malawi's pilot cash transfer programme showed the value of the approach, though the results were not sufficient to support a scale-up over time. In 2012, 3ie funded a second impact evaluation. The evidence informed four changes in the programme design, including its eventual scale-up across all districts in Malawi:

- The size of the cash transfer increased;
- The programme was scaled up to all districts;
- Children's case management improved;
- Linkages to other programmes improved, notably referrals to social services.

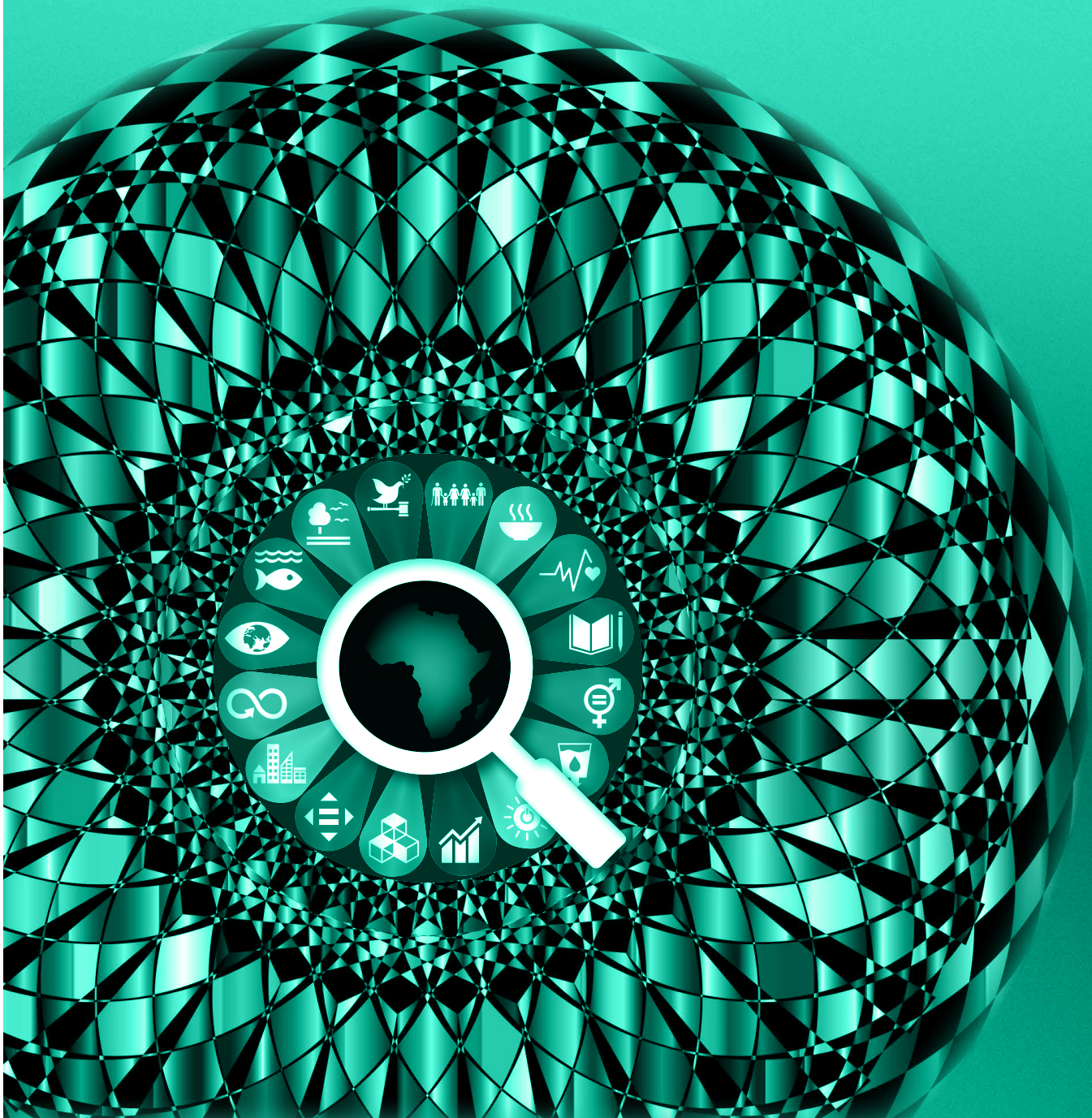
**Source:** Jha Kingra, K. and B. Leach, 'How Evidence Helped Address Ultra Poverty in Malawi', 3ie evidence use brief series, 3ie, New Dehli, 2019, www.3ieimpact.org/evidence-hub/publications/evidence-use/how-evidence-helped-address-ultra-poverty-malawi accessed 21 August 2019.



The VNRs mostly provide statistics on changes in performance that occur over time (e.g., children's school attendance, maternal mortality rates, etc.). The VNR guidelines are not clear on the use of evaluative evidence and the only reference to evaluation is rather generic (monitoring and evaluation systems or frameworks are often mentioned). In practice, the lack of concrete evaluation guidelines means that it is difficult for countries to understand how to produce a rigorous explanation of why any of the expected/desired SDG-related changes are happening, and whether they are happening because of either specific policies or programmes (i.e., the change can be attributed to these policies) or some unrelated effects (e.g., changes in weather, exchange rates or prices of crops). This means that either VNRs are not reporting the findings of past evaluation work or that relevant evaluations, which may help understand and improve national performance against the different SDG targets, are not happening. For example, the Kenya 2017 VNR (p. 14) states that "The Ministry of Devolution and Planning is responsible for the overall management and coordination of the implementation, monitoring and reporting of SDGs process in Kenya." However, evaluation is not mentioned and, although ministries are deemed responsible for the monitoring and evaluation of their specific sectors, no sectoral evaluation is mentioned. On the contrary, Benin, which has an established national evaluation system, produced a VNR that specifically mentioned four evaluations, including an evaluation of the poverty reduction strategy (the PRSP). Section 5 provides more detailed guidance on impact evaluation.

# PART B

## GUIDANCE ON EVALUATIONS AND EVALUATIVE TOOLS



# 4 DIFFERENT TYPES OF EVALUATION



Part B provides more detailed guidance on some of the practical elements of evaluation. Table 4 shows indicative costs and durations for different evaluative processes.

### Table 4: Indicative costs and timings of evaluation processes

| EVALUATION TYPE | INDICATIVE COST (US$) | INDICATIVE DURATION (TO APPROVAL OF REPORT) |
|---|---|---|
| **Diagnostic review** | US$ 100,000–200,000, depending on how much primary data needs to be collected | 8–18 months |
| **Implementation/ process evaluation** | | |
| **Outcome evaluation** | US$ 100,000–250,000, depending on the complexity of the evaluation | |
| **Impact evaluation** | Over US$ 500,000 if significant survey work needed | 3–5 years if baseline/endline needed; 1 year if theory-based approach used |
| **RAPID VERSIONS OF THE ABOVE** | | |
| **Rapid evaluation**[29] | US$ 40,000 | 2–3 months |
| **Evaluative workshop** | US$ 7,000 plus the cost of workshop and participant travel | 1 month, including preparing for workshop, writing up |

**Source:** Developed by the authors with inputs from Dr. Laurenz Langer, University of Johannesburg.

[29] A cost-benefit exercise could fit into this if done alone.

The rest of this section takes each evaluation type (including rapid processes) and summarizes key factors. Each is also considered according to how the Organisation for Economic Co-operation and Development - Development Assistance Committee evaluative criteria are relevant.[30] These criteria are as follows (abbreviated, and with the word 'aid' removed as they equally apply to government interventions):

- **Relevance**: The extent to which the activity is suited to the priorities and policies of the target group, recipient and donor. To what extent are the objectives of the programme still valid?
- **Effectiveness**: A measure of the extent to which an intervention attains its objectives. To what extent were the objectives achieved or likely to be achieved? What were the major factors influencing the achievement or non-achievement of the objectives?
- **Efficiency**: Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs. It is an economic term that signifies that the intervention uses the least costly resources possible to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs to see whether the most efficient process has been adopted. Were activities cost-efficient? Were objectives achieved on time? Was the programme or project implemented in the most efficient way compared to alternatives?
- **Impact**: The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the intervention on the local social, economic, environmental and other development indicators. What has happened as a result of the programme or project? What real difference has the activity made to the beneficiaries? How many people have been affected?
- **Sustainability**: Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after initial (e.g., donor) funding has been withdrawn. Projects need to be socially, environmentally, institutionally and financially sustainable. To what extent did the benefits of a programme or project continue after donor funding ceased? What were the major factors that influenced the achievement or non-achievement of sustainability of the programme or project?

A general source of information is BetterEvaluation. For example, this webpage gives some guidance on how to select the evaluation type.[31]

[30] Organisation for Economic Co-operation and Development, 'DAC Criteria for Evaluating Development Assistance', www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm, accessed 13 November 2019.

[31] See https://betterevaluation.org/en/start_here/decide_which_method, accessed 13 November 2019.

## 4.1 TYPES OF EVALUATION

### 4.1.1 Implementation/process evaluation

| | |
|---|---|
| **Key question** | Is the intervention working or not working and why? This seeks to improve the efficiency and efficacy of operational processes and will evaluate whether an intervention's operational mechanisms support achievement of the objectives or not and why. It looks at activities, outputs and outcomes, the use of resources and causal links. |
| **Examples of evaluation purpose** | The purpose of an implementation evaluation is to understand how a policy, plan or programme is working and how it can be strengthened. An example of an evaluation purpose is:<br>• To assess whether the business process services scheme is being implemented as specified, to explain the performance and propose how to strengthen it. |
| **Typical evaluative questions** | • What theory of change best describes how the intervention was intended to operate?<br>• How is the theory of change working in practice? Are there variations of implementation at different sites or by different service deliverers?<br>• What factors influence the way the programme is implemented?<br>• What have been the start-up and continuing costs of implementation?<br>• What are the strengths and weaknesses of the programme (from the point of view of staff, clients, experts)?<br>• How might the programme be strengthened/upscaled/replicated? |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, sustainability |
| **Implemented when** | An implementation evaluation can happen any time after the programme has been implemented, as a standalone evaluation, as part of a series of evaluations or as one component of an impact or economic evaluation. Typically, it would be undertaken after two to three years. |
| **Methods** | An implementation evaluation can be rapid, primarily using secondary data, or in-depth with extensive field work. It has a strong formative element (learning how to improve) using process/qualitative tools, but will likely also involve quantitative elements (e.g., assessing uptake of services). It may involve case studies. It is likely to involve a review of the theory of change and the design. |
| **Typical tools** | • Document review/literature review of comparative interventions<br>• Development and review of theory of change and log frame<br>• Key informant interviews and document review for case studies<br>• Field work at different sites to assess implementation in practice<br>• Workshopping/interviews with stakeholders<br>• Reviewing budget and expenditure data |
| **Key issues to consider** | • To what extent do you want to get a picture of the likely achievement of outcomes?<br>• Having reasonable monitoring data is important |
| **Key resources** | • DPME Guideline on Implementation Evaluation No. 2.2.12[32]<br>• https://betterevaluation.org/en/plan/approach/case_study |

[32] Republic of South Africa Department for Planning, Monitoring and Evaluation, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.12, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.12%20Implementation%20Evaluation%2014%2003%2020.pdf, accessed 13 November 2019.

### 4.1.2 Outcome evaluation

| | |
|---|---|
| **Key question** | Are the intended outcomes of the intervention being achieved? Are there unintended outcomes and why? An outcome evaluation can be used to see whether the intervention is working and should be continued or not, and if any potential modification is needed. |
| **Examples of evaluation purpose** | The purpose of an outcome evaluation is to understand the effectiveness of an intervention – the changes in outcomes intended to improve the well-being of the target population. An example of an evaluation purpose is:<br>• To assess whether an HIV programme is effective in changing the sexual behaviour of X target group and any unintended outcomes and to understand why and how to improve programme effectiveness. |
| **Typical evaluative questions** | • Is the intervention effective in achieving its intended outcomes, for example:<br>- Were medical providers who received intensive training in sexually transmitted diseases more likely to effectively counsel, screen and treat patients than those who did not?<br>- Did the implementation of sexually transmitted disease counselling in community-based organizations result in changes in knowledge, attitudes and skills among the members of the target population?<br>• Are there any unintended outcomes and what are the implications of these?<br>• How do differences in implementation affect intervention outcomes?<br>• Is the intervention more effective for some participants than for others?<br>• How do the outcomes compare with an alternative intervention?<br>• Does the intervention seem sustainable? |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, sustainability |
| **Implemented when** | After three or more years, when outcomes should start being evident |
| **Method** | This can be qualitative using theory-based evaluation or case studies and quantitative comparing different approaches or assessing the state of achievement of intermediate outcomes where monitoring data is not available (e.g., the proportion of children eating free school meals). Evaluation ideally tests whether the outcomes of the theory of change are being achieved. It can also use counterfactual methods such as impact evaluation. |
| **Typical tools** | • Document review/literature review of comparative interventions<br>• Development and review of theory of change and log frame<br>• Appreciative inquiry<br>• Outcome mapping with stakeholders<br>• Community surveys/participatory tools with stakeholders<br>• Most significant change<br>• Workshopping/interviews with stakeholders<br>• Counterfactual methods (e.g., randomized control trials, comparison group analysis, regression analyses, logically created or expert constructed counterfactuals) |
| **Key issues to consider** | • Often undertaken with implementation evaluation to explain the results<br>• If longer-term outcomes are looked for, then it is very close to impact evaluation |
| **Key resources** | Some similar resources to impact evaluations (see section 4.1.3):<br>• https://www.k4health.org/toolkits/measuring-success/outcome-and-impact-evaluation<br>• https://blogs.lse.ac.uk/impactofsocialsciences/2014/10/27/impact-vs-outcome-harding/<br>• https://www.outcomemapping.ca/outcome-mapping-practitioner-guide |

### 4.1.3 Impact evaluation

| | |
|---|---|
| **Key question** | While outcome/effectiveness evaluation assesses the progress towards the outcomes that the intervention aims to achieve, impact evaluation assesses the intervention's ability to achieve (and be responsible for) its ultimate goals. The key question is: Is the intervention having an impact and why? |
| **Examples of evaluation purpose** | The purpose of an impact evaluation is to measure the changes in well-being of the target population (impact) attributable to a specific intervention. An example of an evaluation purpose is:<br>• To assess whether the child support grant is leading to sustained impacts on the levels of education and malnutrition of children. |
| **Typical evaluative questions** | • What are the intended and unintended impacts on the target group?<br>• Can you attribute changes to the intervention or are they due to other factors?<br>• How do differences in implementation affect intervention outcomes?<br>• Is the intervention more effective for some participants than for others?<br>• How do the results compare with an alternative intervention to achieve the same objective?<br>• Are the results sustainable? |
| **Key evaluation criteria** | Relevance, effectiveness, impact, sustainability |
| **Implemented when** | The evaluation should ideally be designed prior to implementation (so that roll-out can be targeted to compare those receiving with those not), with a baseline implemented early and impact checked at key stages (e.g. after 5 to 10 years, possibly after the intervention is complete). |
| **Method** | This can be qualitative using theory-based evaluation or case studies. An impact evaluation design can be experimental or quasi-experimental, whereby it compares groups receiving/not receiving the intervention or receiving it later, etc. This type ideally tests whether the outcomes and impacts of the theory of change are being achieved. |
| **Typical tools** | • Document review/literature review of comparative interventions<br>• Development and review of theory of change and log frame<br>• Appreciative inquiry<br>• Community surveys/participatory tools with stakeholders<br>• Most significant change<br>• Workshopping/interviews with stakeholders<br>• Counterfactual methods (e.g., randomized control trials, comparison group analysis, regression analyses, logically created or expert constructed counterfactuals) |
| **Key issues to consider** | • Counterfactual methods are quite expensive and take a long time but give a very precise quantitative picture of impact<br>• Implement with implementation evaluation to be able to explain the results |

| | |
|---|---|
| **Key resources** | • https://betterevaluation.org/en/themes/impact_evaluation<br>• Elliott Stern, 'Impact Evaluation - A guide for commissioners and managers'[33]<br>• DPME Guideline on Impact Evaluation No. 2.2.13[34]<br>• There are now key partners that can support impact evaluation including 3ie (www.3ieimpact.org) or the Abdul Latif Jameel Poverty Action Lab (www.povertyactionlab.org/africa).[35] |

## 4.1.4 Economic evaluation

| | |
|---|---|
| **Key question** | How do the costs compare to the benefits of the intervention? This type of evaluation can be used to inform high-level officials on the extent to which the intervention should be continued or not, and if any potential modifications are needed. |
| **Examples of evaluation purpose** | The purpose of economic evaluation is the comparative analysis of alternative courses of action in terms of both their costs and outcomes. An example of an evaluation purpose is:<br>• To assess the costs in relation to the benefits of early childhood development centres, compared to home-based provision. |
| **Typical evaluative questions** | • What have been the outcomes and impacts of the programmes/alternative options, in economic/social/environmental/institutional terms, and were these intended or unintended?<br>• What have been the costs of running each of the programmes/alternatives?<br>• How cost-effective has each programme been?<br>• Should one of these be expanded/selected rather than the other? |
| **Key evaluation criteria** | Effectiveness, efficiency, impact, sustainability |
| **Implemented when** | This can be undertaken at any stage, for example, diagnostic (comparing which option to take forward to design), during implementation (what seem to be the cost benefits of different services) or part of impact (having assessed the impact of certain interventions, how does the cost benefit of each compare). |
| **Method** | The main difference between different forms of economic evaluations is how they itemize and value effects:<br>• Cost-benefit analysis involves measuring costs and benefits in common units, usually monetary.<br>• Cost-effectiveness analysis measures benefits in natural units such as life years saved or improvements in functional status. |

[33] Stern, Elliot, 'Impact Evaluation: A guide for commissioners and managers', May 2015, https://assets.publishing.service.gov.uk/media/57a0896de5274a31e000009c/60899_Impact_Evaluation_Guide_0515.pdf, accessed 13 November 2019.

[34] Republic of South Africa Department for Planning, Monitoring and Evaluation, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.13, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.13%20Impact%20Evaluation%2014%2003%2020%20%20(3).pdf, accessed 13 November 2019.

[35] The key contact for the Abdul Latif Jameel Poverty Action Lab in Africa is Laura Poswell at lposwell@povertyactionlab.org. 3ie is supporting a regional project in West Africa to support impact evaluation (West African Capacity Building and Impact Evaluation Program), and the key contact is Deo-Gracias Houndolo at dhoundolo@3ieimpact.org.

◀ 4.1.4 Economic evaluation (cont'd)

| | |
|---|---|
| **Typical tools** | • Document review/literature review of comparative interventions<br>• Development and review of theory of change and log frame<br>• Appreciative inquiry<br>• Community surveys/participatory tools with stakeholders<br>• Most significant change<br>• Workshopping/interviews with stakeholders<br>• Counterfactual methods (e.g., randomized control trials, comparison group analysis, regression analyses, logically created or expert constructed counterfactuals) |
| **Key issues to consider** | • Availability of detailed cost data<br>• Availability of good estimates of effects/benefits<br>• Need to consider non-economic as well as economic effects |
| **Key resources** | • https://betterevaluation.org/en/evaluation-options/CostBenefitAnalysis<br>• DPME Guideline on Economic Evaluation No. 2.2.15[36] |

## 4.1.5 Diagnostic review

| | |
|---|---|
| **Key question** | What is the core problem in a sector/area? What are the underlying root causes? How do different options compare for addressing this? |
| **Examples of review purpose** | The purpose of a diagnostic review is to understand the problem facing a sector/area and what interventions are needed to address it. An example of a purpose of such a review is:<br>• To assess the current situation of malnutrition and the root cause of the problem, to what extent current interventions are addressing this and what needs to be done to strengthen state response. |
| **Typical evaluative questions** | • What is the extent of the problem (e.g., malnutrition) in the country and for whom? How does this compare to other countries?<br>• What are the root causes of the problem?<br>• What programmes are being undertaken to address the problem?<br>• What is recognized as the best practice in this area?<br>• How does the state's response compare to other countries?<br>• What options should be considered for addressing the problem, bearing in mind the interventions currently in place? What are the advantages and disadvantages of these?<br>• What intervention(s) are recommended |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, impact, sustainability |
| **Implemented when** | This can be done prior to designing an intervention, or after a period of time to get an overview of the sector/area. |
| **Method** | This can involve qualitative methods such as case studies, as well as quantitative methods, for example to assess the situation. |

[36] Republic of South Africa Department for Planning, Monitoring and Evaluation, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.15, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.15%20Economic%20Guideline%20%2014%2003%2020%20docx%20%20-%20Copy.pdf, accessed 13 November 2019.

◀ 4.1.5 Diagnostic review (cont'd)

| | |
|---|---|
| **Typical tools** | • Document review/literature review of situation and comparative interventions<br>• Root cause analysis (e.g., problem trees, usually with stakeholders)<br>• Participatory tools with stakeholders<br>• Most significant change<br>• Workshopping/interviews/focus groups with stakeholders<br>• Options analysis |
| **Key issues to consider** | • Availability of detailed data or this may need to be collected<br>• Ensuring there is a good selection of stakeholders to participate (e.g., in root cause analysis) to get a good understanding |
| **Key resources** | • DPME Guideline on Diagnostic Evaluation No. 2.2.10[37]<br>• www.roma.odi.org/diagnosing_the_problem.html |

## 4.2 RAPID EVALUATIVE METHODS

### 4.2.1 Rapid evaluations

| | |
|---|---|
| **Key question** | This is likely to be a version of an implementation or economic evaluation, or possibly a diagnostic review preceding planning. It can be used:<br>• To support learning and improvement as innovations are tested out in real-world settings where practitioners and policymakers are hungry for evidence;<br>• Because innovations and contexts change during implementation – ongoing feedback of findings ensures relevance and usefulness;<br>• To generate evidence to sustain innovations beyond initial pilots (and where timescales for decisions are often tight);<br>• To provide timely information about potential (or not) for scale up/roll-out across the wider system;<br>• Practical considerations (e.g., funders may not be able to commit resources for longer-term assessments).[38] |
| **Examples of rapid evaluation purpose** | This will depend on which of the above types of evaluation this is a rapid version of (e.g., diagnostic review, implementation, outcome, impact). |
| **Typical evaluative questions** | As mentioned in sections 4.1.1–4.1.5 |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, impact, sustainability |
| **Implemented when** | This can be done prior to designing an intervention or when a particular crisis or problem emerges that requires a rapid response, where the trade-off of limited rigour is acceptable. |

[37] Republic of South Africa Department for Planning, Monitoring and Evaluation, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.10, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.10%20Diagnostic%20Evaluation%2014%2003%2020.pdf, accessed 13 November 2019.

[38] Smith, Judith, 'Rapid Evaluation in Healthcare: Opportunities, need and challenges', www.nuffieldtrust.org.uk/files/2019-02/judith-smith.pdf , accessed 13 November 2019.

◀ **4.2.1** Rapid evaluations (cont'd)

| | |
|---|---|
| **Method** | This can include qualitative approaches such as case studies and quantitative approaches primarily using existing data or through an electronic survey with a limited number of respondents expected. This can be a formative exercise and depending on the priority can be done internally. Lessons from experience are more successful when distilled with the client.[39] |
| **Typical tools** | • Document review/literature review of comparative interventions<br>• Development and review of theory of change and log frame<br>• Appreciative inquiry/participatory tools with stakeholders<br>• Workshopping/interviews/focus groups with stakeholders<br>• Electronic survey (e.g., of participants of training) |
| **Key issues to consider** | • How critical for credibility is it that the evaluation is seen to be rigorous. If the priority is time, consider rapid evaluation.<br>• Rapid evaluation needs a more experienced team. Is this available?<br>• Is there good monitoring or administrative data? That can save a lot of time<br>• Do you have good estimates of effects/benefits? Causal data is difficult to verify rapidly. |
| **Key resources** | • http://www.pointk.org/resources/files/TIG_6_Rapid_Eval.pdf<br>• Miles McNall and Pennie G. Foster-Fishman, 'Methods of Rapid Evaluation, Assessment and Appraisal'[40] |

## 4.2.2 Evaluative workshop

| | |
|---|---|
| **Key question** | This is likely to be a process for evaluative thinking around implementation, a diagnostic review preceding planning, or to assess outcomes in a participatory process. |
| **Examples of workshop purpose** | This will depend on which of the above types of evaluation is being used. An example of a real workshop purpose is:<br>• By the end of the workshop, the performance of the National Youth Policy to date since its inception in May 2015 has been assessed and necessary steps to strengthen it have been identified. |
| **Typica evaluative questions** | • This can follow the types mentioned in 4.1.1–4.1.5. Questions could be:<br>• What is the core problem and what are the root causes (develop a problem tree)?<br>• What is the theory of change for X and to what extent is it working?<br>• How has X performed against targets at the outcome and output levels and who has benefited/not?<br>• What are the lessons from X?<br>• What recommendations are there for strengthening X going forward?<br>• How can the institutionalization of X be strengthened? |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, impact, sustainability |

[39] For example, Battye, Fraser and Steph Cartwright, 'Advancing Rapid Evaluation', 29 January 2019, www.nuffieldtrust.org.uk/files/2019-02/fraser-battye.pdf, accessed 13 November 2019.

[40] McNall, Miles and Pennie G. Foster-Fishman, 'Methods of Rapid Evaluation, Assessment and Appraisal', American Journal of Evaluation, vol. 28, no. 2, June 2009, pp. 151-168, www.researchgate.net/profile/Pennie_Foster-Fishman/publication/249773524_Methods_of_Rapid_Evaluation_Assessment_and_Appraisal/links/561fc8c508ae93a5c9243298.pdf, accessed 13 November 2019.

◀ 4.2.2 Evaluative workshop (cont'd)

| | |
|---|---|
| **Implemented when** | This can be done prior to designing an intervention or when a particular crisis or problem emerges that requires a rapid response, where a very rapid response is needed |
| **Method** | The key method is a participatory workshop. The key to the method is the careful development of the process of the event, the quality of facilitation, careful selection of participants to give a 360 degree view of the issues to be addressed and the quality of preparation. |
| **Typical tools** | • Group work<br>• Exercises such as problem tree, theory of change, participatory learning for action tools<br>• Mixed groups of stakeholders<br>• Potential for preparation prior to bringing material to review |
| **Key issues to consider** | • It is essential to bring people from the field level to the policy level who understand the unit of analysis. This should include external stakeholders, such as different client groups, good field staff and good mid-level managers.<br>• Preparation by a good facilitator typically takes the same time as the length of the workshop. Do you also want them to write up a report?<br>• Establish a facilitation team between the facilitator and some key staff from the agency to support the process.<br>• What level of authorization is needed to give urgency and meaning to the event (e.g., opening by the Director General). |
| **Key resources** | • https://betterevaluation.org/en/evaluation-options/rapid_outcome_assessment<br>• http://www.roma.odi.org/diagnosing_the_problem.html |

## 4.2.3 Annual review models

| | |
|---|---|
| **Key question** | This is likely to be an annual process for reflecting on an intervention's progress and the changes needed. |
| **Review purpose: an example** | To review progress on the intervention, lessons and changes needed over the next period |
| **Typical evaluative questions** | Questions could be:<br>• What has been the progress against the outputs?<br>• Is there evidence of progress towards the outcomes?<br>• Does it look like the theory of change is working?<br>• What evidence of efficiency is there?<br>• Does it look like the intervention is sustainable?<br>• What are the key lessons this year?<br>• What changes should be implemented next year? |
| **Key evaluation criteria** | Relevance, effectiveness, efficiency, impact, sustainability |
| **Implemented when** | This can be implemented on an annual basis during the implementation of an intervention. |
| **Method** | The key method is a review by a small team of external evaluators (could be funders), ideally with the partner agencies. |

▶

◀ **4.2.3** Annual review models (cont'd)

| | |
|---|---|
| **Typical tools** | • Preparation of report by the project team to use as a basis<br>• Document review<br>• Small number of interviews of key stakeholders<br>• Use of monitoring data<br>• Meeting with the project team |
| **Key issues to consider** | • Good to have this as a joint exercise, so that the main stakeholders (including funders) can learn and benefit from it<br>• Decide whether some independence is needed (e.g., to have one or more independent members of the team)<br>• Very helpful if a prior report is done by the project/programme team against the key questions, providing sources of evidence. |
| **Key resources** | United Kingdom Department for International Development guidance for annual reviews and project completion reviews[41] |

## 4.3 DECIDING WHAT TYPE OF EVALUATIVE PROCESS TO UNDERTAKE

Table 4 shows indicative costs and durations for different evaluative processes to consider when deciding what sort of process to undertake. Key factors include the budget and time available and the internal skills to undertake these processes. It is also possible to do evaluations quite cheaply provided that, for example, Master's degree students are employed to undertake it. However, this requires careful supervision. Other factors to take into account include data availability (this is needed generally and particularly for some types of impact evaluation).

High-level policymakers in country are usually interested in outcomes (per cent of children in school) or impacts (levels of maternal mortality or that by introducing the child support grant, malnutrition dropped by X per cent). This means that in order to measure the effectiveness and levels of achievement of any given policy and programme, an outcome or impact evaluation will be necessary. However, implementation evaluations may be sufficient for determining whether an intervention is on course to achieve the intended outcomes.

The SDGs are usually at the level of long-term outcome or impact. For example, taking Goal 4 – ensure inclusive and equitable quality education and promote lifelong learning opportunities for all – some targets are at the outcome level (e.g., by 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing states), while others are at the impact level (e.g., by 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy). This means that understanding levels of achievement of the SDGs will also require outcome or impact evaluations, while implementation evaluations may be sufficient to indicate whether these outcomes are likely to be achieved.

It would be good to compare this with other African countries. There are rapid evaluation options that are not nearly as rigorous (e.g., field work may be minimal), but where it is possible to get a good indication of root causes or progress with implementation. Twende Mbele (www.twendembele.org) is currently supporting the testing of methodologies around the use of rapid evaluation.[42]

[41] Department for International Development, 'Reviewing and Scoring Projects', DFID practice paper, November 2011, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/67344/HTN-Reviewing-Scoring-Projects.pdf, accessed 13 November 2019.

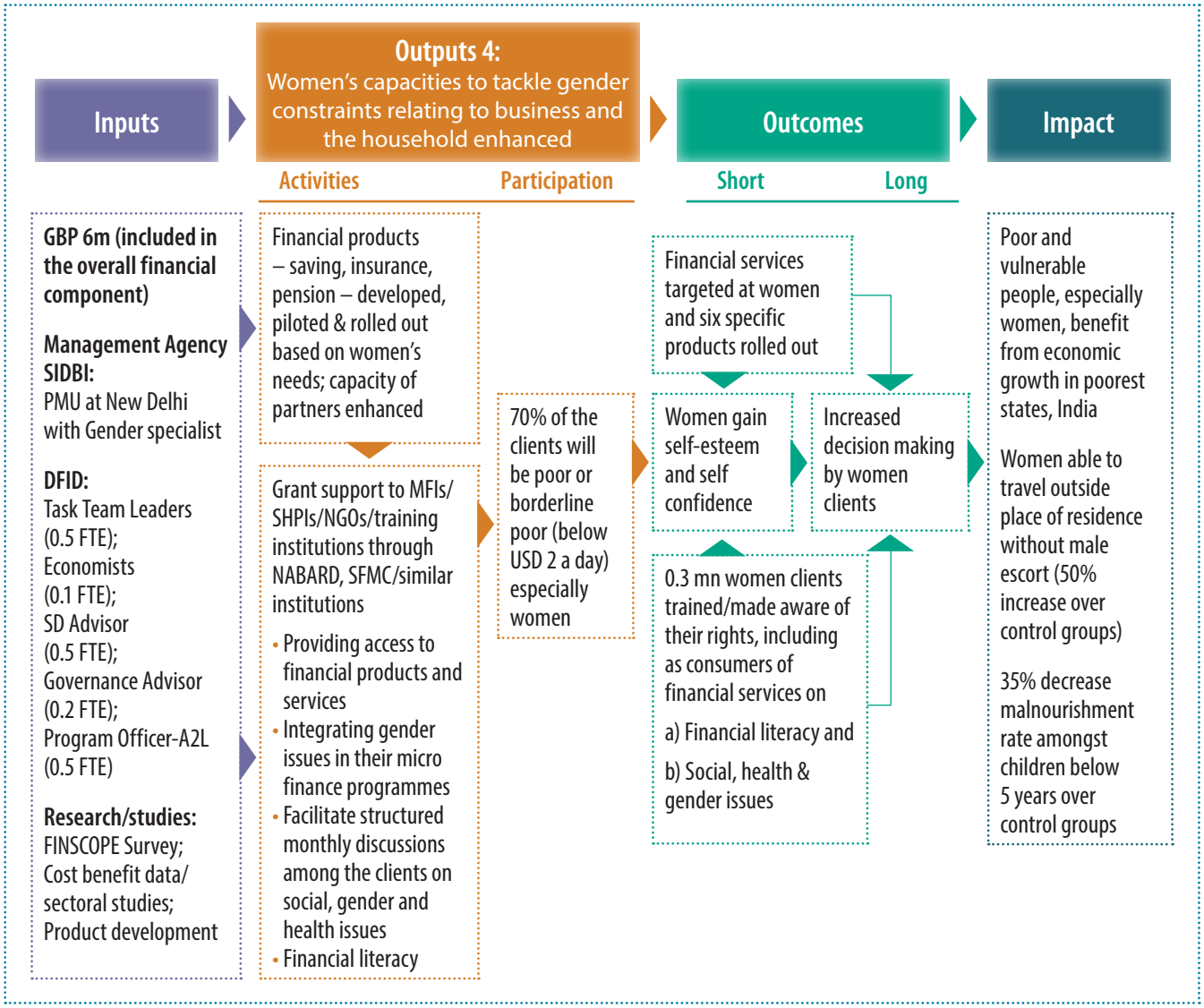
[42] The key contact is Cara Waller, Programme Manager, Twende Mbele (cara.waller@wits.ac.za).

# 5 OTHER EVALUATIVE TOOLS

## 5.1 IMPORTANCE AND USE OF A THEORY OF CHANGE

A theory of change provides a way of understanding how an intervention is supposed to work, in a diagrammatic format. Having a developed theory of change makes it much easier to plan an evaluation, even when programme documentation is inadequate. Figure 4 shows a useful definition of the hierarchy of changes using the output/outcome/impact definitions, as often used for logical frameworks.[43] Figure 5 shows an example of a theory of change.

**Figure 5: Example of a theory of change for the India Poorest States Inclusive Growth Programme[44]**




[43] A useful resource differentiating between these levels is Simister, Nigel, 'Outputs, Outcomes and Impact', INTRAC, 2015, www.intrac.org/wpcms/wp-content/uploads/2016/06/Monitoring-and-Evaluation-Series-Outcomes-Outputs-and-Impact-7.pdf accessed 13 November 2019.

[44] Drawn from Vogel, Isabel and Zoe Stephenson, 'Appendix 3: Examples of theories of change', DFID EVD, July 2012, https://assets.publishing.service.gov.uk/media/57a08a66ed915d622c000703/Appendix_3_ToC_Examples.pdf, accessed 13 November 2019. This source provides a range of examples of theories of change.

## Guide for developing a theory of change

| | |
|---|---|
| **Objective** | The theory of change, sometimes referred to as programme theory or programme logic, is a way of tracking how the desired outcomes will be achieved, the indicators at different stages and the assumptions required to achieve the outcomes. |
| **Implemented when** | This is ideally implemented as the first stage of the planning/design quadrant of the policy/programme cycle. In evaluations, it is a very useful tool to start with to define, with stakeholders, what the intervention aimed to do, which the evaluation can test, and then propose a revised theory of change. |
| **Typical steps** | 1. This should be preceded by an adequate diagnosis, identifying the root causes, comparing interventions and then selecting the proposed intervention. It is helped considerably if there is a problem tree or fishbone.<br>2. Organize a meeting of a group of stakeholders who understand the intervention, preferably including frontline staff and/or clients. About one day will be needed for this, and it should be facilitated by someone with experience with theories of change.<br>3. Introduce the concept of theories of change.<br>4. If you are designing a new intervention, then go downwards as indicated below. If you are developing a theory of change for an existing programme, for example for evaluating it, then it is often easier to go upwards, starting with activities, to outputs, outcomes and impacts.<br>5. Brainstorm the **outcomes** on cards and organize them in a causal logic of at least two to three levels of outcome and impact (*see Figure 1*). Do not write one word, write four to five words to make what you are saying clear.<br>6. Brainstorm the **outputs** to achieve the outcomes and put these on cards. Outputs you have control over, while outcomes you influence. Arrange these in a flow diagram to show the linkages leading to the outcomes.<br>7. Brainstorm the **activities** needed to achieve the outputs.<br>8. Identify the **assumptions** needed for each card/level to reach the next level (e.g., if output A happens and assumption X holds true then immediate outcome B should result).<br>9. If appropriate, identify key **indicators** and put these on cards (you may also move from the theory of change to develop a **log frame**, in which case you will develop indicators in that process).<br>10. Check the logic going downwards (in order to achieve impact A, then wider outcomes B, C and D are needed, and assumptions E, F and G need to hold). In order to achieve outcome B, etc.<br>11. Check the logic going upwards (if I do these activities and these assumptions hold, then I should achieve this output).<br>12. Share the theory of change with others to get feedback.<br>13. Identify appropriate **evaluative questions** from the theory of change that you would want to ask. |
| **Resources needed** | • Wall with large sheets of paper (at least four flip chart sheets)<br>• Cards in different colours<br>• Felt pens (ideally black/blue) – at least one per team<br>• Sticky stuff/Prestik to stick cards on the paper |
| **Sources of further information** | • The BetterEvaluation website is a useful resource of information on evaluation and provides a good set of resources on theories of change.[45]<br>• Funnell, S. C. and P. J. Rogers, 'Purposeful Program Theory: Effective use of theories of change and logic models', Jossey-Bass/Wiley, San Francisco, 2011. |

[45] See www.betterevaluation.org/en/search/site/theory%20of%20change and http://betterevaluation.org/plan/define/develop_logic_model , accessed 13 November 2019.

## 5.2 CREATING AN EVALUATION REPOSITORY

| | |
|---|---|
| **Background** | Evaluations have been carried out in most countries, often funded by donors or non-governmental organizations. For example, a search in Uganda identified over 500 evaluations. This is a tremendous resource that can be used immediately to inform planning, VNR reports, etc. |
| **Objective** | The objective is to identify the existing evaluations that have been conducted and use these to inform policy and decision-making as well as the VNR. |
| **Implemented when** | This is ideally implemented as the first stage of establishing an evidence base from evaluation. |
| **Typical steps** | Identify evaluations already conducted:<br>1. Consult CLEAR AA's African Evaluation Database46 which has over 2,600 evaluations in 12 African countries (Botswana, Ethiopia, Ghana, Kenya, Namibia, Nigeria, Rwanda, South Africa, Uganda, the United Republic of Tanzania, Zambia and Zimbabwe).<br>2. Consult the 3ie impact evaluation repository.47 See Figure 6 for an individual record that can then be accessed to have a page or so of detail and links to reports and sometimes the data.<br>3. Identify an information specialist/librarian (e.g., at a university) to undertake systematic database searching for "evaluation" and the country. Note some evaluations may also use the words "review" or "assessment".<br>4. Write to all ministries, donors, major non-governmental organizations and evaluation service providers asking them for evaluations conducted and copies of the reports.<br>5. Use consultants' proposals to identify evaluations already conducted.<br>6. Follow up on these to identify at least one paper or one report that can be used to interrogate findings.<br>7. Create a database for these reports or papers that is accessible online and searchable by keyword/sector.<br>8. Communicate with planners in the different sectors to indicate that this resource is available to assist in diagnosis and planning. |
| **Resources needed** | • Access to an information specialist who themselves has access to online databases and papers<br>• Resources to develop an online database<br>• Allocation of staff time to follow up on this, or a service provider |
| **Tips** | When requesting ministries/donors to provide evaluations, have a letter from a very senior person, as they may be reluctant to provide these. |
| **Sources of further information** | • Benin's repository.[48]<br>• South Africa's evaluation repository.[49] |

[46] See https://crest2.sun.ac.za/african_evaluation_db/default/african_eval_db_01, accessed 13 November 2019.

[47] See www.3ieimpact.org/evidence-hub/impact-evaluation-repository, accessed 13 November 2019.

[48] See www.presidence.bj/evaluation-politiques-publiques, accessed 13 November 2019.

[49] See https://evaluations.dpme.gov.za/evaluations.aspx, accessed 13 November 2019.

## Figure 6: Example of an evaluation listed in the 3ie repository

Does a multi-faceted market-based approach to food crops stimulate food security and agricultural development in Tanzania?

Impact evaluation | 3ie | 2019

**Publication type:** 3ie final grantee report
**Author:** Marrit van den Berg, Ian Levely
**Sector:** Agriculture and Rural Development, Agricultural Extension, Agro-Industry & Marketing

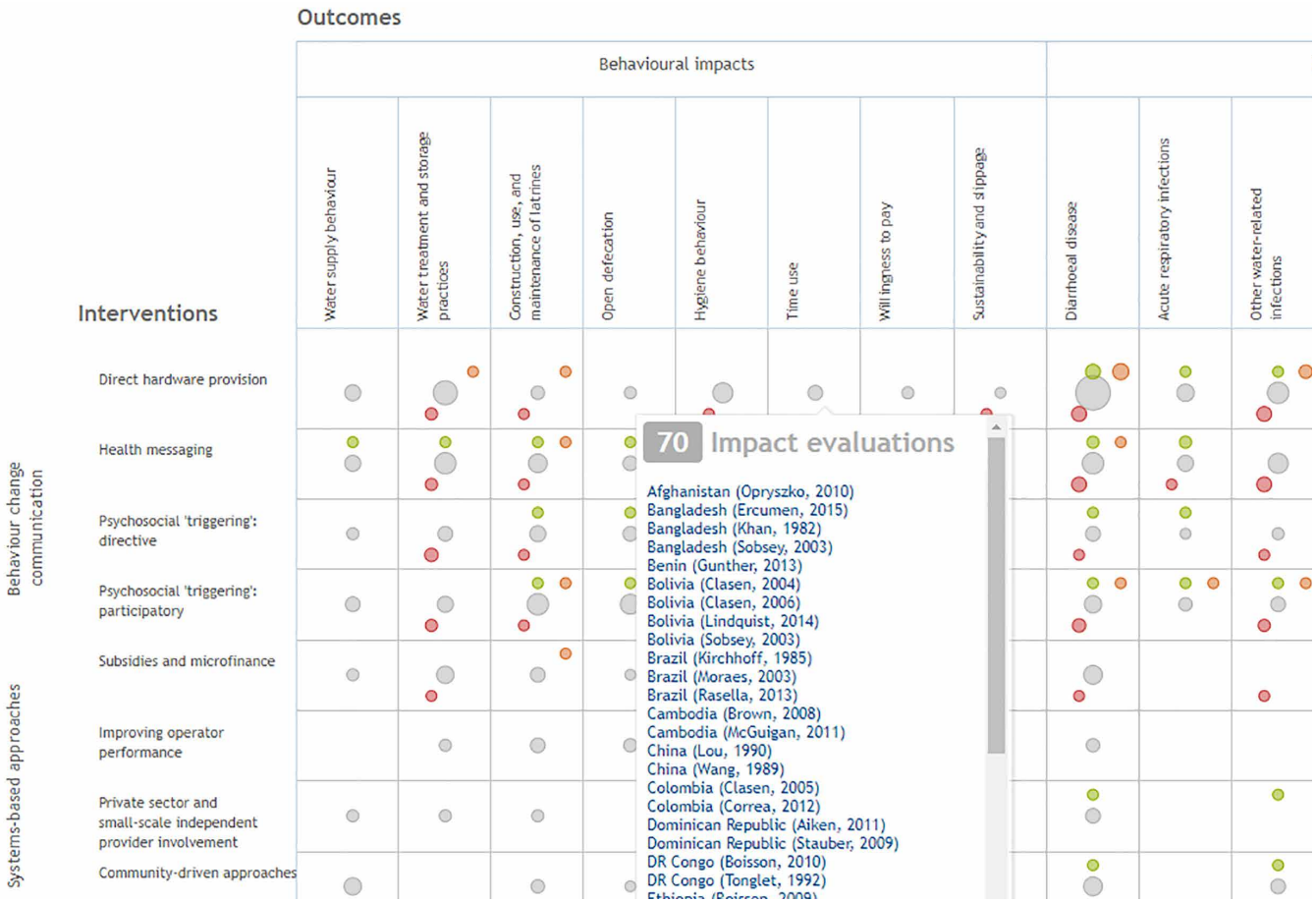
**Source:** International Initiative for Impact Evaluation.

## 5.3 DEVELOP AN EVIDENCE (GAP) MAP USING EVALUATIONS

| | |
|---|---|
| **Background** | In addition to searching for evaluations generally, these can be used to populate an evidence map, sometimes referred to as evidence gap map, where these evaluations are categorized (e.g., by goals in a national development plan and usually by type of intervention). They present a visual overview of existing and ongoing studies or reviews in a sector or sub-sector in terms of the types of programmes evaluated and the outcomes measured. This evidence is mapped onto this framework, graphically highlighting the gaps, where few or no impact evaluations or systematic reviews exist and where there is a concentration of impact evaluations but no recent high-quality systematic review (*see Figure 7*). An evidence map like this is being created for Uganda, and several evidence maps have been generated by DPME in South Africa, as well as other departments. |
| **Objective** | The objective is to identify existing evaluations conducted to inform policy and decision-making, as well as the VNR, and identify the gaps where new evaluations are needed. |
| **Implemented when** | This is ideally implemented as the first stage of establishing an evidence base from evaluation. |
| **Typical steps** | 1. Use the steps indicated for the evaluation repository to identify evaluations.<br>2. Screen these evaluations for relevance to narrow down the search.<br>3. Screen those that are relevant for quality (those already screened, for example on the 3ie database, can be used directly).<br>4. Write summaries of the relevant studies.<br>5. Create an evidence map electronic platform to hold these studies, with each cell being an intersection of an outcome and an intervention, holding the number of studies and the ability to link to those studies (*see Figure 7*).<br>6. Communicate with planners in the different sectors to indicate that this resource is available to assist in diagnosis and planning and run a training session to tell them how to use it. |

| | |
|---|---|
| **Resources needed** | • Access to an information specialist who themselves has access to online databases and papers<br>• Resources to develop an online evidence map, or use the 3ie evidence map platform<br>• Allocation of resources to commission someone to support this, and for your staff to be part of the process |
| **Tips** | • Use the 3ie platform to create an evidence map.<br>• The University of Johannesburg African Centre for Evidence specializes in evidence maps and runs training. |
| **Sources of further information** | • 3ie's evidence maps[50]<br>• Contact Birte Snilstveit of 3ie (bsnilstveit@3ieimpact.org)<br>• Contact Ruth Stewart of the African Centre for Evidence (ruths@uj.ac.za)<br>• Contact Harsha Dayal (harsha@dpme.gov.za) |

## Figure 7: Example of an evidence map from the 3ie evidence map repository on water, sanitation and hygiene




**Source:** International Initiative for Impact Evaluation, 2018, <http://gapmaps.3ieimpact.org/evidence-maps/water-sanitation-and-hygiene-wash-evidence-gap-map-2018-update>, accessed 13 November 2019.

[50] See www.3ieimpact.org/evidence-hub/evidence-gap-maps, accessed 13 November 2019.

## 5.4 SYNTHESIZING EVIDENCE FROM EXISTING EVALUATIONS

| | |
|---|---|
| **Background** | Once sets of evaluations have been identified, this existing evidence can be synthesized to show what the existing evidence is saying. This could be from the repository or evidence map already suggested, or if the country has done several evaluations in a particular sector. For example, South Africa has done seven evaluations and an evidence map in the human settlements sector, and a synthesis has been done from these. Conducting an evaluation synthesis is important for several reasons. First, the synthesized findings from several evaluations can help decision-makers make better judgments by bringing together available evidence that can be critically appraised (for quality) within the decision-making timeframe. Second, evaluation synthesis makes use of existing evaluation and research about programmes and issues in service delivery, so it draws on more data than a single evaluation. This brings a broad perspective. Third, evaluation synthesis integrates evaluation findings, establishes an accessible knowledge base and identifies knowledge gaps or needs for a specific topic or theme, which can then guide future evaluations. |
| **Objective** | The purpose of evaluation synthesis is to systematically distil and integrate data from a number of evaluations and other sources of evidence to draw more informed conclusions about a given question or to inform planning and policy-making and potentially the VNRs. |
| **Implemented when** | This is ideally implemented once an initial evidence base of evaluations has been created. This could be simply after identifying evaluations from the African Evaluation Database and 3ie databases. |
| **Typical steps** | 1. Use the steps indicated for the evaluation repository to identify relevant evaluations for the particular question you have, or for the sector.<br>2. Obtain the reports, or at least papers on the findings.<br>3. Either commission a service provider or undertake a synthesis yourselves drawing out the common findings and lessons from these for your question/sector.<br>4. Develop recommendations from this body of evidence. |
| **Resources needed** | • Access to a set of evaluations, either from a repository or evidence map<br>• Allocation of resources to commission someone to support the synthesis, and for your staff to be part of the process |
| **Tips** | • At a minimum, search the African Evaluation Database and 3ie databases to get evaluations to draw on for a qualitative synthesis.<br>• It is possible to move to a rigorous process of synthesis, either a rapid synthesis or a full systematic review. This follows strict protocols so the work can be replicated. |
| **Sources of further information** | • https://betterevaluation.org/en/rainbow_framework/synthesise/synthesise_across_evaluations<br>• DPME Guideline on Evaluation Synthesis No. 2.2.15[51]<br>• Makerere University has centres supporting systematic reviews as well as rapid synthesis. Contact Dr. Ekwaro Obuko at the Africa Centre for Systematic Reviews and Knowledge Translation, Makerere University, (ekwaro@gmail.com)<br>• 3ie has a repository of systematic reviews[52]<br>• To follow up on systematic reviews, contact Marie Gaarder (mgaarder@3ieimpact.org) or Ruth Stewart at the African Centre for Evidence (ruths@uj.ac.za). |

[51] Republic of South Africa Department for Planning, Monitoring and Evaluation, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.15, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.15%20Economic%20Guideline%20%2014%2003%2020%20docx%20%20-%20Copy.pdf, accessed 13 November 2019.

[52] See www.3ieimpact.org/evidence-hub/systematic-review-repository, accessed 13 November 2019.

## Table 5: Forms of evaluation synthesis, their nature and time to complete

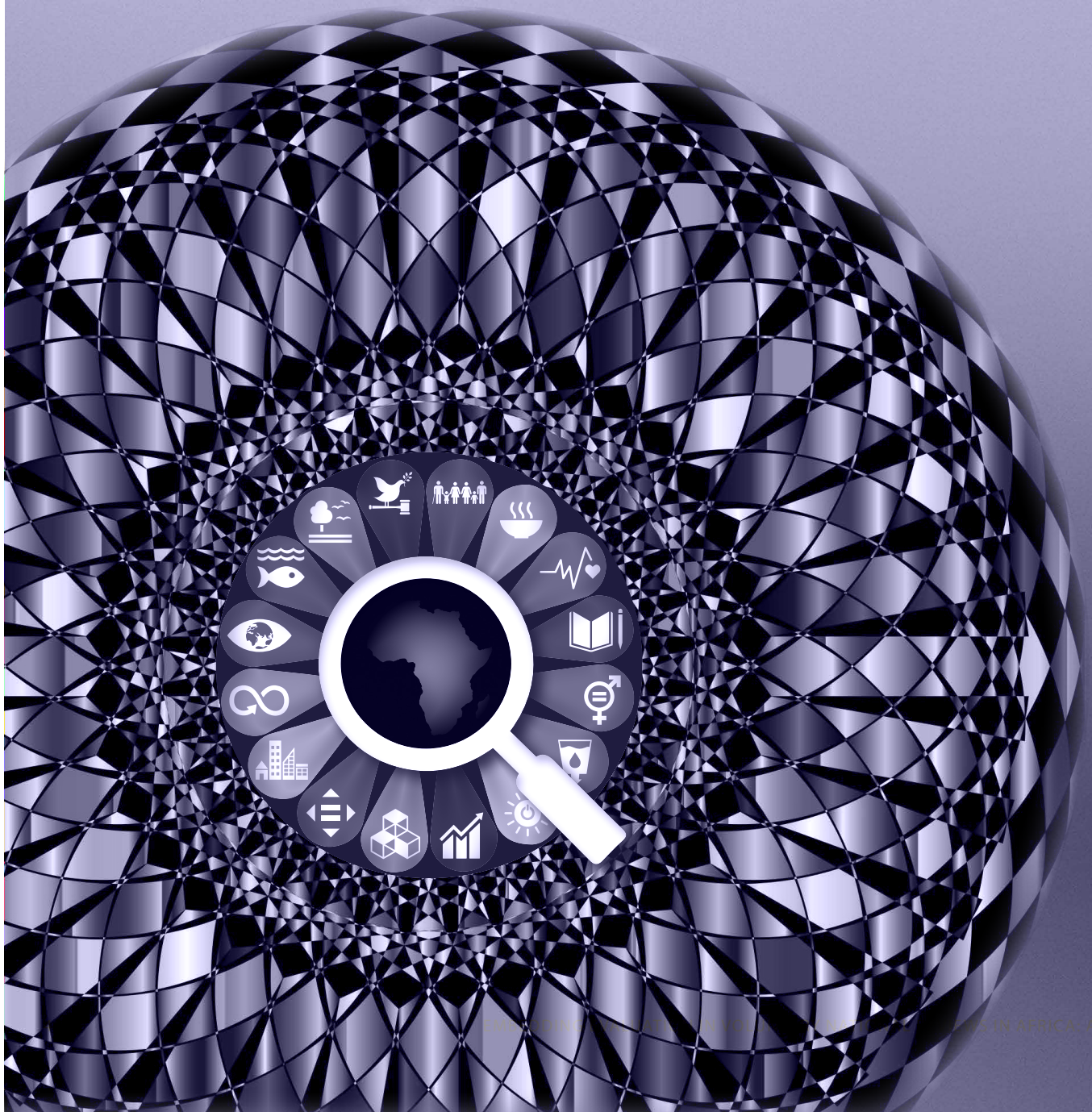
| TYPE OF SYNTHESIS | NATURE | TIME REQUIRED TO COMPLETE |
|---|---|---|
| Traditional literature review (sometimes called narrative review)[53] | Narrative, selective review (not systematic); collates relevant studies and draws conclusions from them. Does not use a systematic method to identify studies. Often also does not have a systematic method for synthesizing or prioritizing studies and dealing with conflicting findings. | 1 week to 2 months |
| Quick scoping review | Quick overview of research undertaken on a (constrained) topic. This could be systematic, but because it is quick, it is unlikely to be comprehensive or detailed. | 1–2 weeks to 2 months |
| Rapid evidence assessment | Systematic search but a quick overview of existing research on a topic. Synthesis of evidence provided by these studies to answer the rapid evidence assessment question. | 2–6 months (quicker than systematic review) |
| Full systematic review | A broad review of existing research on a topic and synthesis of the evidence provided by these studies to answer the review question. | 8–12 months minimum |
| Multi-arm systematic review | Full map and synthesis of different types of evidence to answer the review question. | 12 months minimum |
| Review of reviews | Same as any of the above methods but only includes reviews of other reviews. | Often quicker than other types of full systematic review |

**Source:** DPME, 'Guideline on Implementation Evaluation', No. 2.2.15.

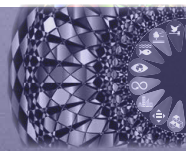
[53] Narrative review can get confused with narrative synthesis which can take place within a systematic review.

# PART C

## GOING BEYOND EVALUATION TO EVALUATION ECOSYSTEMS



# 6 EVALUATION ECOSYSTEMS



## 6.1 WHAT IS AN EVALUATION SYSTEM OR ECOSYSTEM?

For evaluation to move beyond an individual study, a number of elements need to be in place, ranging from management of the system, to resourcing it, ensuring quality, providing capacity development, etc. The list provided in Table 6 can eventually be covered. In countries that have established national evaluation systems, such as Benin, South Africa and Uganda, many of these elements are in place. Goldman et al. (2018) describe these elements.[54] Figure 8 illustrates one way of characterizing the elements of the system. Table 6 adapts this and describes some elements of the evaluation system.

A special edition of the African Evaluation Journal describes the process of establishing South Africa's national evaluation system.[55] For evaluation to become systematized, these different elements need to be implemented. Although this is not a quick project, using the experience of Benin, South Africa and Uganda suggests that it will take 5 to 10 years to establish all of the elements of the ecosystem. The key is developing and encouraging the development of a national evaluation association that can foster professional evaluation. The parent body for these associations is the African Evaluation Association (AFREA).[56]

### Figure 8: Elements of an evaluation system

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Policy | The Evaluation Plan | M vs. E | Autonomy and Impartiality | Feedback | Alignment planning and budgeting | |
| 2 | Methodology | Selection of indicators | Selection criteria | Priority setting | Causality chain | Methodologies used | Data collection |
| 3 | Organisation | Coordination and oversight | Statistical office | Line ministries | Decentralised levels | Link with projects | |
| 4 | Capacity | Problem acknowledged | Capacity building plan | | | | |
| 5 | Participation of actors outside government | Parliament | Civil society | Donors | | | |
| 6 | Quality | Effective use of M&E | Internal usage of M&E findings | | | | |

**Source:** Holvoet, N. & renard, R., 2007, 'Monitoring and evaluation under the PRSP: Solid rock or quick sand?', *Evaluation and Program Planning* 30, 66-81, https://doi.org/10.1016/j.evalprogplan.2006.09.002

[54] Goldman, et al. 'The Emergence of Government Evaluation Systems in Africa: The case of Benin, Uganda and South Africa', Afr. Eval. J. vol. 6, no. 11, 2018, https://doi.org/10.4102/aej.v6i1.253, accessed 13 November 2019.

[55] These papers can be accessed at African Evaluation Journal, https://aejonline.org/index.php/aej/issue/view/6, accessed 13 November 2019.

[56] See African Evaluation Association, www.afrea.org, accessed 13 November 2019.

## Table 6: Key elements of the evaluation ecosystem

| ELEMENT | WHAT IS NEEDED |
|---|---|
| **Governance** | Key governance structures (e.g., cabinet meetings) require evidence |
| | Evaluations of organization's work tabled at parliamentary committees and shared with wider public |
| | Politicians and senior managers read evidence that is provided in an accessible way |
| **Policy and/or legislation** | A policy guides the development of the system. This can either be a monitoring and evaluation policy (as in Uganda) or an evaluation policy (as in South Africa). This could be embedded in legislation. Benin is currently developing legislation. |
| **Plans** | Plans for evidence needed |
| | Improvement plans in place and monitored following evaluations |
| | Organizational plans require demonstration of the use of evidence in diagnosis and design |
| **Management** | Management systems require the use of evidence and the implementation of improvement plans |
| | Leadership and culture incentivizes learning, innovation and use of evidence |
| | Structures/meetings require specific evidence to be discussed and tabled |
| | Community of practice emerging where the evaluation profession manages itself, represented by a national evaluation association (Voluntary Organizations of Professional Evaluation (VOPE)) |
| **Participation of stakeholders** | The level of participation of different stakeholders within government, academia and civil society. VOPE is an important stakeholder. |
| **Funding** | Funding for evidence production and sharing |
| **Human resources** | Structures include evaluation, research and data positions with the appropriate skills and these posts are funded |
| | Development of standard job descriptions, including generation, management and use of evidence |
| | Performance agreements include generation and use of evidence |
| **Capacity development** | Competencies developed related to generation and use of evidence for government staff, as well as monitoring and evaluation and research professionals, ideally linked to a national evaluation association (VOPE) |
| | Curricula reflect policy needs and competences for evidence generation and use |
| | Formal courses, short courses and mentoring systems in evidence generation and use (evaluation, evidence maps, systematic reviews etc.) |
| **Quality assurance** | Standards for evidence production developed, including evaluation, evidence maps, etc. |
| **Brokering and communication** | Units to play knowledge broker role in place in government and linked to those outside |
| | Evaluation and research evidence stored and accessible (e.g., evaluation repository) |
| | Systematic communication of results from monitoring and evaluation and research |
| | Media share results from evidence generation and use |
| | Evidence produced in accessible forms |

**Source:** Developed by the authors.

## 6.2 PRACTICAL WAYS TO BUILD THE EVALUATION SYSTEM USING THE VOLUNTARY NATIONAL REVIEW PROCESS

The VNR process can assist the development of an evaluation system in a number of ways:

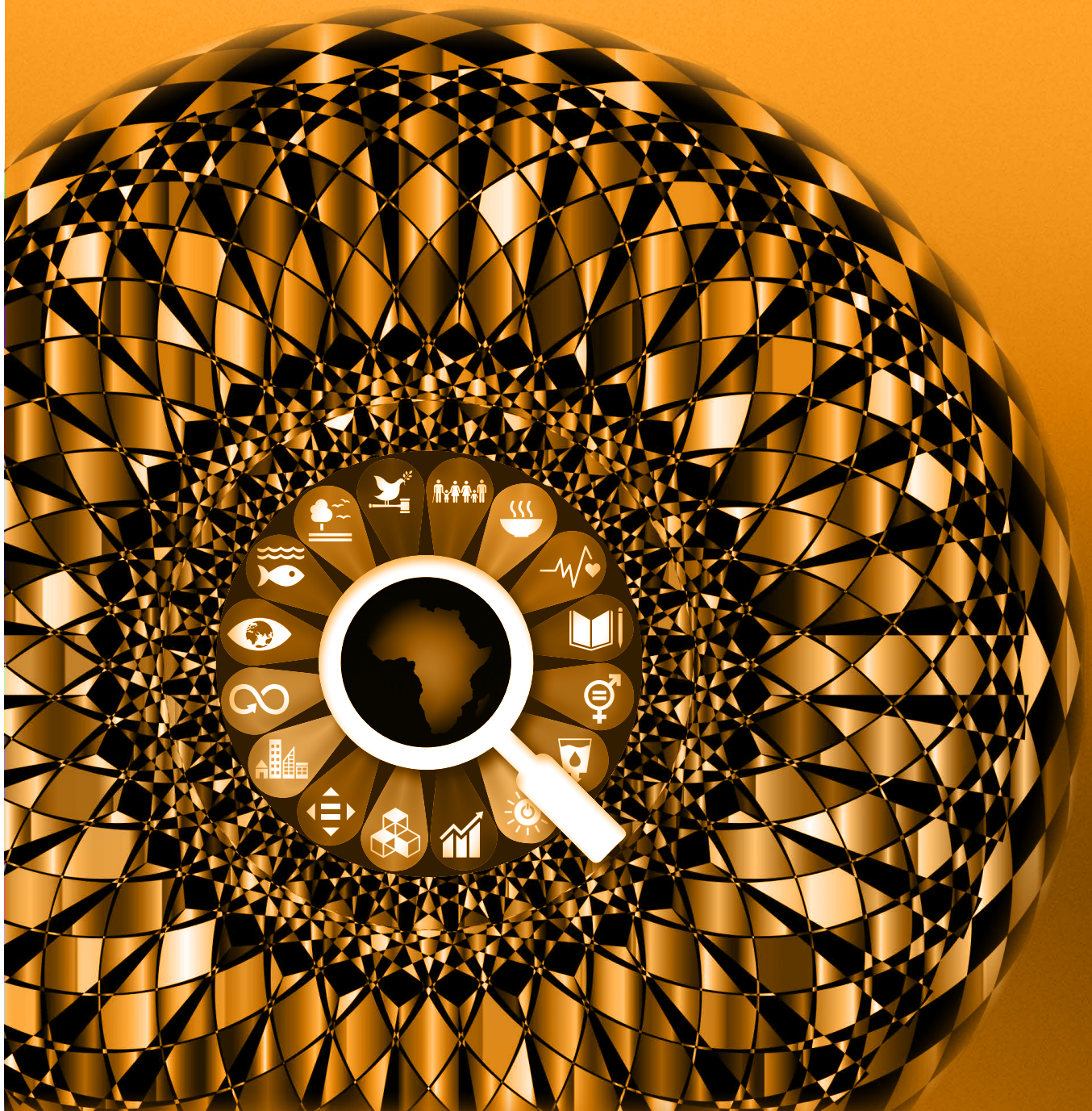
- Encouraging the use of evaluation evidence in programmes and policies that are reported on in the VNR;
- Using the VNR to stimulate the drawing together of evidence, existing evidence or new evaluations or evaluative processes;
- Building a coalition of agencies to support the use of evaluation evidence, including UNICEF and other United Nations agencies;
- Engaging UNICEF and other development partners to contribute with pilot evaluations, evaluative workshops or an inventory of existing evaluations as part of the VNR process, which can help to establish an evaluation system;
- Using the VNR to raise the profile of the need for evidence in country;
- Encouraging in-country capacity to synthesize or undertake evaluations.

A pilot evaluation provides a good way to establish a system. Davids et al. (2015) describe how this was done in South Africa.[57] This pilot was used to test out different elements of the system, ranging from how to do terms of reference, the inception phase, management response and improvement plans. As each stage was tested and found to be relevant, guidelines were developed. In parallel, the systems for national evaluation plans, evaluation standards and competencies were developed.

[57] Davids, M. et al., 'The Pilot Evaluation for the National Evaluation System in South Africa: A diagnostic review of early childhood development', Afr. Eval. J. vol. 3, no. 1, 2015, https://doi.org/10.4102/aej.v3i1.141.

# PART D

## APPLYING THIS IN THE VOLUNTARY NATIONAL REVIEWS



# 7 APPLYING THIS IN THE 2020 –2021 VOLUNTARY NATIONAL REVIEWS

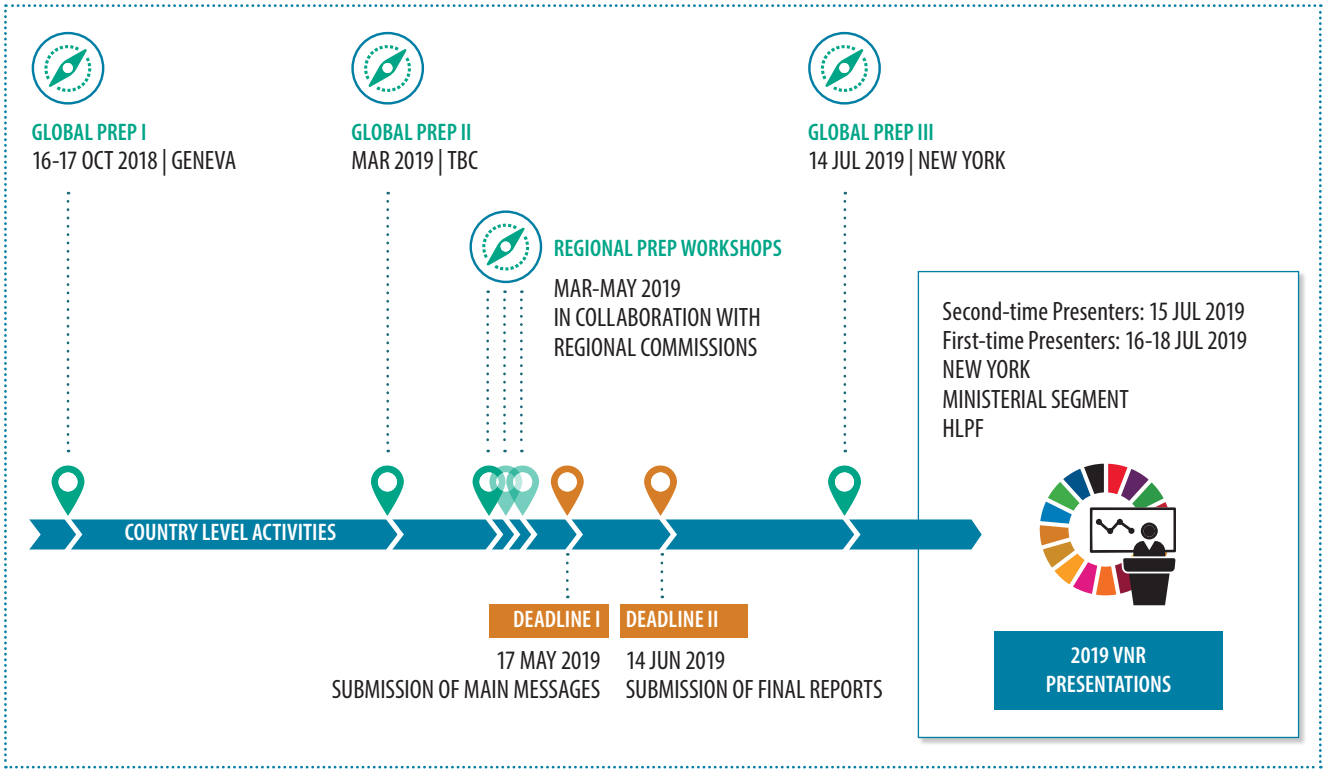
## 7.1 APPLYING THIS IN THE 2020 VOLUNTARY NATIONAL REVIEW

The VNRs will have to be tabled at the High-Level Political Forum at the United Nations in July 2020, typically in a 15-minute presentation. The timescale for preparation for this in 2019 is shown in Figure 9.

If a similar timescale is followed for the 2020 VNR, there are around 13 months for preparation. What are some of the activities that countries can undertake in the 12 to 13 months available to them for the 2020 VNRs? Some examples include:

- **Audit** what evaluations have been undertaken in the country by the government, by donors and by civil society organizations, and obtain them where possible. For those relevant to the 2019 themes, synthesize the lessons (*see section 5.2*). Other relevant sources include country performance audits for the SDGs undertaken by supreme audit institutions in some countries.[58]
- Use the carefully prepared stakeholder workshops as **evaluative workshops** to identify areas where further analytical work is needed on the topics for the 2020 VNR (*see section 4.2.2*).

### Figure 9: Timeline for 2019 Voluntary National Reviews




[58] See the following guide to preparing these reviews on preparedness to implement the SDGs in the Middle East and North Africa: INTOSAI, 'A Practical Guide to Government SDG Preparedness Reviews: Based on the experiences and reflections of seven supreme audit institutions', July 2018, www.intosai.org/fileadmin/downloads/focus_areas/SDGs_and_SAIs/SDGs_Netherlands_EN.pdf. Algeria, Morocco and Tunisia have implemented this.

- Develop **theories of change** for key programmes and use these to check that the designs and indicators are appropriate (*see section 5.1*).
- Undertake **rapid evaluations** where there are significant gaps around specific policies or programmes (*see section 4.2.1*).
- Commission **synthesis** work on key gaps where broader knowledge of what does or does not work is needed (*see sections 5.3–5.4*).
- Meet with the national monitoring and evaluation champions in the executive, Parliament (e.g. Parliamentary Research Service) and national monitoring and evaluation associations, and **discuss how country-led evaluations can be undertaken** going forward, or even how national evaluation systems can be promoted, drawing from examples from Benin, South Africa and Uganda. There could be some support from Twende Mbele[59] or from the African Parliamentary Network on Development Evaluation (APNODE).
- In **West Africa, discuss with the West African Capacity Building and Impact Evaluation (WACIE) Program** how they could support the undertaking of specific evaluations or the establishment of evaluation systems.
- Discuss with **UNICEF, CLEAR-AA** and **CLEAR Francophone Africa** the possibility of supporting evaluation systems.
- Discuss with UNICEF the undertaking of **pilot evaluations** that could pilot both the undertaking of evaluations and the development of evaluation systems in country and help to build local capacity around evaluation.

## Table 7: Possible sequence of activities for 2020 VNR, drawing on evaluative evidence

| ACTIVITY | 2019 | | | | | | 2020 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | JUL | AUG | SEP | OCT | NOV | DEC | JAN | FEB | MAR | APR | MAY | JUN | JUL |
| Stakeholders meet to agree on priority areas for reporting and develop an action plan (including a schedule like this) | | | | | | | | | | | | | |
| Commission evidence maps on these areas, including searching for existing evaluations | | | | | | | | | | | | | |
| Decide on where rapid evaluative activities should be undertaken in gaps | | | | | | | | | | | | | |
| Commission rapid evaluations | | | | | | | | | | | | | |
| Commission syntheses drawing from reports identified in the evidence map | | | | | | | | | | | | | |
| Workshop with stakeholders emerging findings and the implications of these | | | | | | | | | | | | | |
| Develop a draft outline report and draw out the key areas of action needed | | | | | | | | | | | | | |
| Workshop with stakeholders | | | | | | | | | | | | | |
| Revise report | | | | | | | | | | | | | |
| Get sign off on report | | | | | | | | | | | | | |
| High-Level Political Forum | | | | | | | | | | | | | |

[59] Contact Cara Waller, Programme Manager, Twende Mbele: cara.waller@wits.ac.za.

## 7.2 PLANNING FOR THE 2021 VOLUNTARY NATIONAL REVIEW CYCLE

The 2020 VNR provides many more options, including those described above. These could include:

- Work closely with stakeholders to develop an **evidence agenda** around key areas of concern, including the topics the country plans to report on to the High-Level Political Forum.
- In particular, this can guide where an overview of existing evidence should be obtained, using **evidence maps** (section 5.3) or evaluation synthesis (section 5.4). This can help inform the **evidence agenda**.
- Where there are gaps, **plan for full evaluations** or research studies, particularly in areas where monitoring is showing there are problems or funding decisions are needed. The priority should be on implementation evaluations or theory-based outcome evaluations, where it is feasible to get results in just over one year that can be fed into policy processes. Some countries with major investments (e.g., in pensions, child support grants, etc.) may wish to undertake some impact evaluations.
- Develop **improvement plans** arising from the evaluations.[60]
- Think about how the development of a **national evaluation system** can strengthen country policies and programmes, and how the results can be used in VNRs.
- Look at **parallel/shadow reports** that are being done by stakeholders in some countries.[61]

### Table 8: Possible sequence of activities for 2021 VNR, drawing on evaluative evidence

| ACTIVITY | 2019 | | 2020 | | | | 2021 | | |
|---|---|---|---|---|---|---|---|---|---|
| | JUL-SEP | OCT-DEC | JAN-MAR | APR-JUN | JUL-SEP | OCT-DEC | JAN-MAR | APR-JUN | JUL |
| Stakeholders meet to agree on priority areas for reporting and develop an action plan (including a schedule like this) | | | | | | | | | |
| Evidence agenda developed and linked to the VNR | | | | | | | | | |
| Commission evidence maps on these areas including searching for existing evaluations | | | | | | | | | |
| Pilot evaluation (including drawing out lessons for wider evaluations and an evaluation system) | Select | Commission | | | Undertake | | Improvement plan | | |

[60] A guide to improvement plans can be found at Republic of South Africa Department for Planning, Monitoring and Evaluation, 'How to Develop an Improvement Plans to Address Evaluation Recommendations', DPME Evaluation Guideline 2.2.6, updated July 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.6%20Improvement%20Plan%2014%2007%2018.pdf.

[61] This note provides guidance for civil society to report on their contribution to the implementation of the 2030 Agenda: Long, Graham, 'How Should Civil Society Stakeholders Report their Contribution to the Implementation of the 2030 Agenda for Sustainable Development?' UN DESA, <https://sustainabledevelopment.un.org/content/documents/18445CSOreporting_paper_revisions_4May.pdf>.

◀ Table 8: Possible sequence of activities for 2021 VNR, drawing on evaluative evidence (cont'd)

| ACTIVITY | 2019 | | 2020 | | | | 2021 | | |
|---|---|---|---|---|---|---|---|---|---|
| | JUL-SEP | OCT-DEC | JAN-MAR | APR-JUN | JUL-SEP | OCT-DEC | JAN-MAR | APR-JUN | JUL |
| As the pilot is being implemented, develop guidelines based on what seems to work (e.g., how to do terms of reference, how to structure evaluation reports, roles of steering committees, etc.) | | | | | | | | | |
| Decide on major evaluations to undertake and use the lessons from above to implement these | | | | | | | | | |
| Undertake major evaluations | | | Commission | | | | | | |
| Decide on where rapid evaluative activities should be undertaken | | | | | | | | | |
| Undertake rapid evaluations | | | | | | | | | |
| Undertake syntheses drawing from reports identified in the evidence map | | | | | | | | | |
| Workshop with stakeholders emerging findings and the implications of these | | | | | | | | | |
| Develop initial reports, validate with stakeholders and develop improvement plans with stakeholders | | | | | | | | | |
| Partners start to implement improvements | | | | | | | | | |
| Develop wider elements of the evaluation system (e.g., competences, standards and courses) | | | | | | | | | |
| Workshop emerging picture with stakeholders based on rapid/full evaluations, syntheses | | | | | | | | | |
| Revise report | | | | | | | | | |
| Get sign off on report | | | | | | | | | |
| High-Level Political Forum | | | | | | | | | |

## 7.3 SUPPORT AVAILABLE

Support may be available from a number of agencies, including regional African organizations such as Twende Mbele, APNODE, the United Nations Economic Commission for Africa, CLEAR and other United Nations agencies such as UNICEF, VOPE, 3ie, etc. This could include technical support, financial support for pilot evaluations and sharing of experience from others. Contacts are below:

- UNICEF offices:
  - UNICEF contact overall: Ada Ocampo, aocampo@unicef.org
  - West Africa: Michele Tarsilla, mtarsilla@unicef.org
  - East Africa: Urs Nagel, unagel@unicef.org
- United Nations Economic Commission for Africa Evaluation Office: Eskedar Nega, Chief of Evaluation, enega@uneca.org
- EvalSDGs: Ada Ocampo (*see above*)
- Twende Mbele: Cara Waller, Programme Manager, Cara.waller@wits.ac.za
- WACIE Program: Deo-Gracias Houndolo, dhoundolo@3ieimpact.org
- AFREA/VOPE: Contact the local VOPE through https://afrea.org/wp-content/cache/all/index.html#
- 3ie: Marie Gaarder, mgaarder@3ieimpact.org
- CLEAR AA: Dugan Fraser, dugan.fraser@wits.ac.za
- CLEAR Francophone Africa: Boubacar Aw, boubacar.aw@cesag.sn
- APNODE: apnode@afdb.org
- Government of Benin: Abdoulaye Gounou, Bureau of Public Policy Evaluation and Government Action Analysis, Presidency, agounou0@gmail.com
- Government of South Africa: Jabu Mathe, DPME, jabu@dpme.gov.za
- Government of Uganda: Tim Lubanga, Office of Prime Minister, tlubanga@gmail.com

## ANNEX 1: REFERENCES


Battye, Fraser and Steph Cartwright, 'Advancing Rapid Evaluation', 29 January 2019, www.nuffieldtrust.org.uk/files/2019-02/fraser-battye.pdf, accessed 13 November 2019.

Cairney, P., 'The Politics of Evidence-Based Policy Making', Palgrave Macmillan, London, 2016.

Davids, M. et al., 'The pilot Evaluation for the National Evaluation System in South Africa: A diagnostic review of early childhood development', Afr. Eval. J. vol. 3, no. 1, 2015, https://doi.org/10.4102/aej.v3i1.141.

Department for International Development, 'Reviewing and Scoring Projects', DFID practice paper, November 2011, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/67344/HTN-Reviewing-Scoring-Projects.pdf, accessed 13 November 2019.

Federal Democratic Republic of Ethiopia, 'Ethiopia 2017 Voluntary National Review on SDGs: Government commitments, national ownership and performance trends', National Planning Commission, Addis Ababa, June 2017.

Goldman, et al., 'Mere compliance or learning – M&E culture in the public service of Benin, Uganda and South Africa', chapter 4 in Goldman, I and Pabari, M (eds) 'Policy Making in Africa: Reflections on Using Evidence', Routledge, forthcoming 2020.

Goldman, et al. 'The Emergence of Government Evaluation Systems in Africa: The case of Benin, Uganda and South Africa', Afr. Eval. J. vol. 6, no. 11, 2018, https://doi.org/10.4102/aej.v6i1.253.

International Initiative for Impact Evaluation, 2018, http://gapmaps.3ieimpact.org/evidence-maps/water-sanitation-and-hygiene-wash-evidence-gap-map-2018-update, accessed 13 November 2019.

International Institute for Environment and Development, 'Effective Evaluation for the Sustainable Development Goals', IIED, 2016, www.iied.org/effective-evaluation-for-sustainable-development-goals, accessed 13 November 2019.

INTOSAI, 'A Practical Guide to Government SDG Preparedness Reviews: Based on the experiences and reflections of seven supreme audit institutions', July 2018, www.intosai.org/fileadmin/downloads/focus_areas/SDGs_and_SAIs/SDGs_Netherlands_EN.pdf, accessed 13 November 2019.

Jha Kingra, K. and B. Leach, 'How Evidence Helped Address Ultra Poverty in Malawi', 3ie evidence use brief series, 3ie, New Dehli, 2019, www.3ieimpact.org/evidence-hub/publications/evidence-use/how-evidence-helped-address-ultra-poverty-malawi, accessed 13 November 2019.

Long, Graham, 'How Should Civil Society Stakeholders Report their Contribution to the Implementation of the 2030 Agenda for Sustainable Development?' UN DESA, https://sustainabledevelopment.un.org/content/documents/18445CSOreporting_paper_revisions_4May.pdf, accessed 13 November 2019.

McNall, Miles and Pennie G. Foster-Fishman, 'Methods of Rapid Evaluation, Assessment and Appraisal', American Journal of Evaluation, vol. 28, no. 2, June 2009, pp. 151-168, www.researchgate.net/profile/Pennie_Foster-Fishman/publication/249773524_Methods_of_Rapid_Evaluation_Assessment_and_Appraisal/links/561fc8c508ae93a5c9243298.pdf.

Organisation for Economic Co-operation and Development, 'DAC Criteria for Evaluating Development Assistance', www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm, accessed 13 November 2019.

Programme to Support Pro-Poor Development in South Africa, 'Evidence-Based Policy-Making (EBPM): Enhancing the use of evidence and knowledge in policy making', PSPPD, July 2011.

Republic of South Africa Department of Planning, Monitoring and Evaluation, 'Evaluations', https://evaluations.dpme.gov.za/evaluations.aspx?keyword=grade%20R, accessed 13 November 2019.

Ibid, 'How to Develop an Improvement Plans to Address Evaluation Recommendations', DPME Evaluation Guideline 2.2.6, updated July 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.6%20Improvement%20Plan%2014%2007%2018.pdf, accessed 13 November 2019.

Ibid, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.10, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.10%20Diagnostic%20Evaluation%2014%2003%2020.pdf, accessed 13 November 2019.

Ibid, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.12, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.12%20Implementation%20Evaluation%2014%2003%2020.pdf, accessed 13 November 2019.

Ibid, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.13, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.13%20Impact%20Evaluation%2014%2003%2020%20%20(3).pdf, accessed 13 November 2019.

Ibid, 'Guideline on Implementation Evaluation', DPME Evaluation Guideline No 2.2.15, 20 March 2014, https://evaluations.dpme.gov.za/images/gallery/Guideline%202.2.15%20Economic%20Guideline%20%2014%2003%2020%20docx%20%20-%20Copy.pdf, accessed 13 November 2019.

Simister, Nigel, 'Outputs, Outcomes and Impact', INTRAC, 2015, www.intrac.org/wpcms/wp-content/uploads/2016/06/Monitoring-and-Evaluation-Series-Outcomes-Outputs-and-Impact-7.pdf, accessed 13 November 2019.

Slifer-Mbacké, L. and S. Barry, 'Etude Nationale sur les Enfants et les Jeunes du System Educatif au Senegal', USAID, 2017.

Smith, Judith, 'Rapid Evaluation in Healthcare: Opportunities, need and challenges', www.nuffieldtrust.org.uk/files/2019-02/judith-smith.pdf, accessed 13 November 2019.

Stern, Elliot, 'Impact Evaluation: A guide for commissioners and managers', May 2015, https://assets.publishing.service.gov.uk/media/57a0896de5274a31e000009c/60899_Impact_Evaluation_Guide_0515.pdf, accessed 13 November 2019.

United Nations, 'High-Level Political Forum on Sustainable Development' Sustainable Development Goals Knowledge Platform, https://sustainabledevelopment.un.org/hlpf, accessed 13 November 2019.

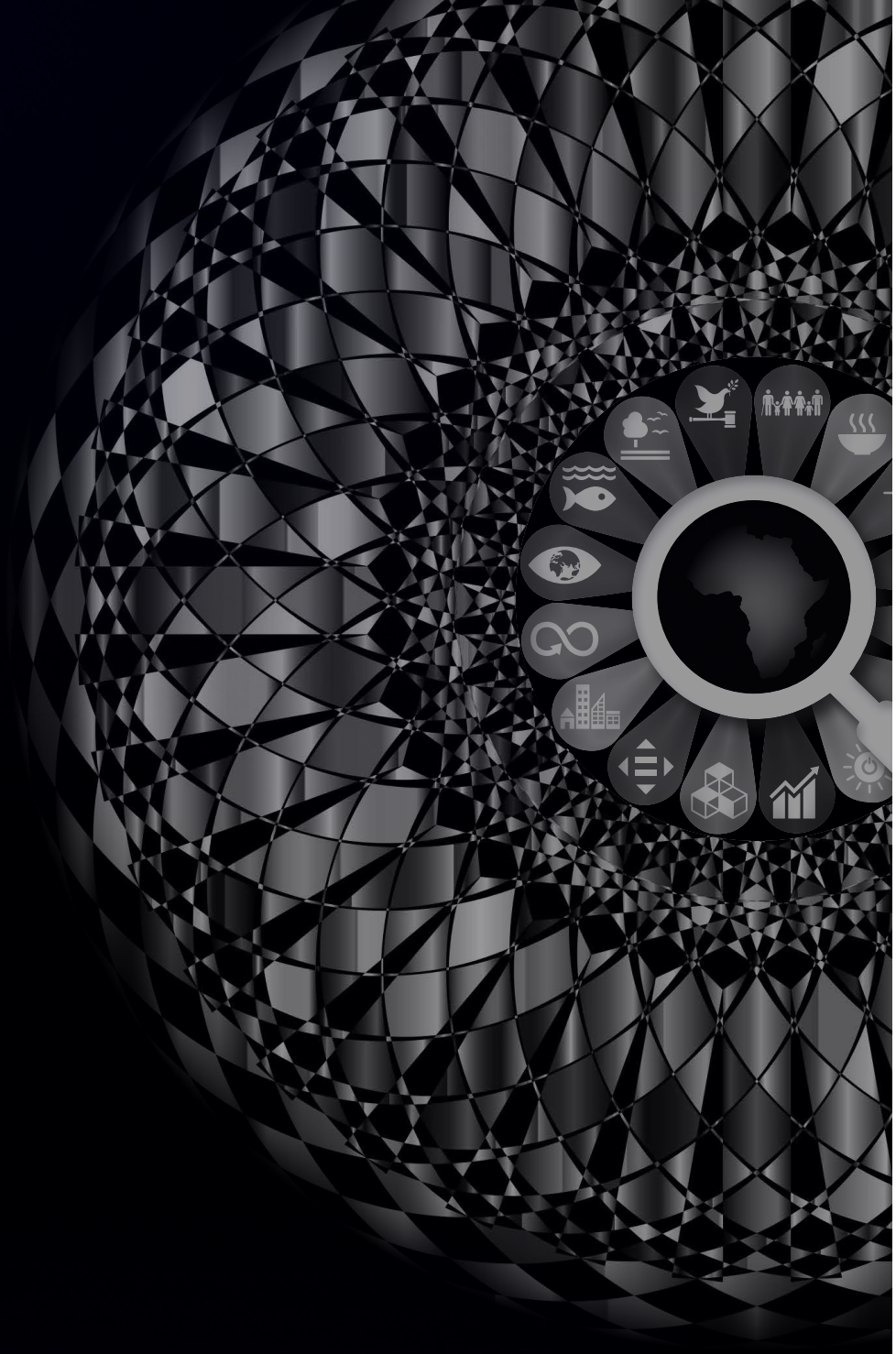
United Nations, 'The Sustainable Development Agenda', Sustainable Development Goals, www.un.org/sustainabledevelopment/development-agenda/, accessed 13 November 2019.

United Nations, 'Transforming our World: The 2030 Agenda for Sustainable Development', Sustainable Development Goals Knowledge Platform, https://sustainabledevelopment.un.org/post2015/transformingourworld, accessed 13 November 2019.

United Nations Department of Economic and Social Affairs, 'Handbook for the Preparation of Voluntary National Reviews', DESA, 2018, https://sustainabledevelopment.un.org/content/documents/20872VNR_hanbook_2019_Edition_v2.pdf, accessed 13 November 2019.

United Nations Department of Economic and Social Affairs, 'Third International Conference on Financing for Development', www.un.org/esa/ffd/ffd3/index.html, accessed 13 November 2019.

Vogel, Isabel and Zoe Stephenson, 'Appendix 3: Examples of theories of change', DFID EVD, July 2012, https://assets.publishing.service.gov.uk/media/57a08a66ed915d622c000703/Appendix_3_ToC_Examples.pdf, accessed 13 November 2019.



clear
Centers for Learning on Evaluation and Results
ANGLOPHONE AFRICA

unicef
for every child