National Evaluation

in the SDG Era:

An Online Self-Assessment Tool for Evaluation Diagnostics and Strategizing

UNDP

Empowered lives.
Resilient nations.

Independent Evaluation Office

United Nations Development Programme

Table of Contents

3

1. INTRODUCTION TO THE ONLINE SELF-ASSESSMENT TOOL FOR NATIONAL EVALUATION DIAGNOSTICS AND STRATEGIZING

A. What is the Self-Assessment Tool for National Evaluation Diagnostics and Strategizing?

The Online Self-Assessment Tool for National Evaluation Diagnostics and Strategizing, developed by the Independent Evaluation Office of UNDP, provides a step-by-step mechanism to assess evaluation capacity, determine needs and establish action points for developing a country's framework for evaluation of national development strategies, including the Sustainable Development Goals (SDGs). The Self-Assessment Tool is aimed at augmenting national evaluation capacities in the SDG era and helping to integrate, strengthen and connect three interlocking elements: national government systems, Agenda 2030 and national evaluation systems and processes.

The Self-Assessment Tool allows a country to use evaluation diagnostics to assess national development strategies and programmes as well as the SDGs. It situates evaluation practices within the context of other public-sector feedback mechanisms and recognizes that an evaluation practice is built over time; it is not a set of outputs.

The Tool enables users to ascertain evaluation bottlenecks and needs for improving evaluation capacities. It also helps to identify needs for critical institutional systems and processes and prepare evaluation frameworks. A brief document summarizing this assessment can be prepared once the self-assessment modules are completed.

The key objectives of this diagnostic process are to:

- Facilitate development or strengthening of a national evaluation framework for the SDGs

5
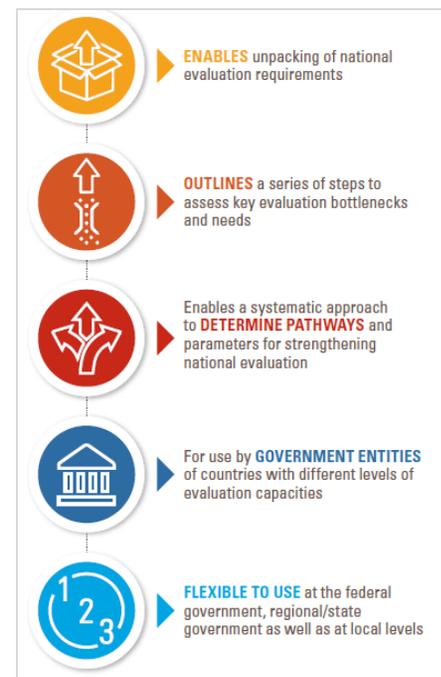
- Inform country-led evaluation processes and systems that are central to the follow-up and review of the progress on the SDGs.

The Self-Assessment Tool provides a flexible and practical framework for a country to assess evaluation capacities of the government as a whole or of specific government entities. Its overall purpose is to enable governments to develop a systematic approach to determining key areas, pathways and parameters for evaluating national development strategies and achievement of the SDGs.

Download full text of the National Evaluation Diagnostics Guidance here http://web.undp.org/evaluation/

B. How does the Tool work?

The Tool consists of a series of questions that enable countries to identify requirement for building evaluation capacity. It helps users to understand the enabling environment for national evaluation systems and institutional capacities in different countries. It also connects national systems to Agenda 2030 and enables integration of the SDGs and sustainable development issues into the evaluation process.

Countries are expected to use the diagnostic results to develop and implement an evaluation action plan that identifies feasible priorities, clear targets, a timeline and required resources for building their national evaluation capacities.

C. Who should use the Tool?

The Self-Assessment Tool is for use by government entities, regardless of the level of their country's evaluation systems, processes and capacities. It can be used by evaluation systems that are mature, evolving or emergent (link to pop-up text). The Tool lends itself to analysis of evaluation needs and bottlenecks and strategies to develop evaluation systems and processes at different levels of government.

The Tool can be used by staff of the central evaluation entity, federal government institutions, sectoral agencies and regional, state and local governments. It can also be used to develop evaluation processes for major national programmes where information on impact is critical. The self-assessment process is driven by the government entity/ies, and the Tool is designed to be simple to use with minimal support from evaluation experts.

DIFFERENT LEVELS OF EVALUATION SYSTEMS AND PROCESSES:

Mature Evaluation Capacity

MAINSTREAM EVALUATION across government institutions

WELL ESTABLISHED systems and processes

SOPHISTICATED information systems to help with complexity of policies and decisions

CHAMPION EVALUATION in global fora and by supporting peers

Evolving Evaluation Capacity

EVALUATE key national priorities and programmes

GENERATE EVIDENCE on key indicators

Demonstration that evaluation use **SUPPORTS GOOD GOVERNANCE**

Systems and processes are **BEING ESTABLISHED**

Emergent Evaluation Capacity

Create **ENABLING ENVIRONMENT** for consultation and critical review of performance

Generate **BASIC EVIDENCE**

OFFICIALS CO-LEAD EVALUATIONS financed by external stakeholders

D. What does the evaluation diagnostic process entail?

The process of carrying out a national evaluation diagnosis matters just as much as its ultimate findings. An accurate and insightful analysis of national evaluation should be perceived as important, credible and useful by relevant stakeholders. Carrying out a diagnostic process therefore consists of much more than gathering evidence and disseminating a report. The diagnostic process includes four phases:

The diagnostic process begins before the collection and analysis of information on evaluation systems and processes, and it does not end with presentation of findings in a document. A carefully managed process adds significant value to the diagnostic exercise. Ideally, the diagnostic process is part of a larger change-management process. It builds on three practices that apply to all phases of the diagnostic process:

- **Fostering ownership**: To facilitate change in national evaluation systems and processes, it is essential to build ownership throughout each phase of the diagnostic process, particularly at the planning and reporting stage. Ownership is built by consulting, involving and informing stakeholders, both through punctual engagements (workshops, presentations) and by formally assigning them responsibilities and powers (as reviewers, committee members, etc.).

- **Leveraging the process for capacity building, raising awareness and sharing information:** The process of conducting a national evaluation diagnostic provides an opportunity to raise stakeholders' awareness of the utility of evaluative practices and to connect them to other stakeholders with shared interests. To do so, it is important to structure consultations not merely as a technical process but one that
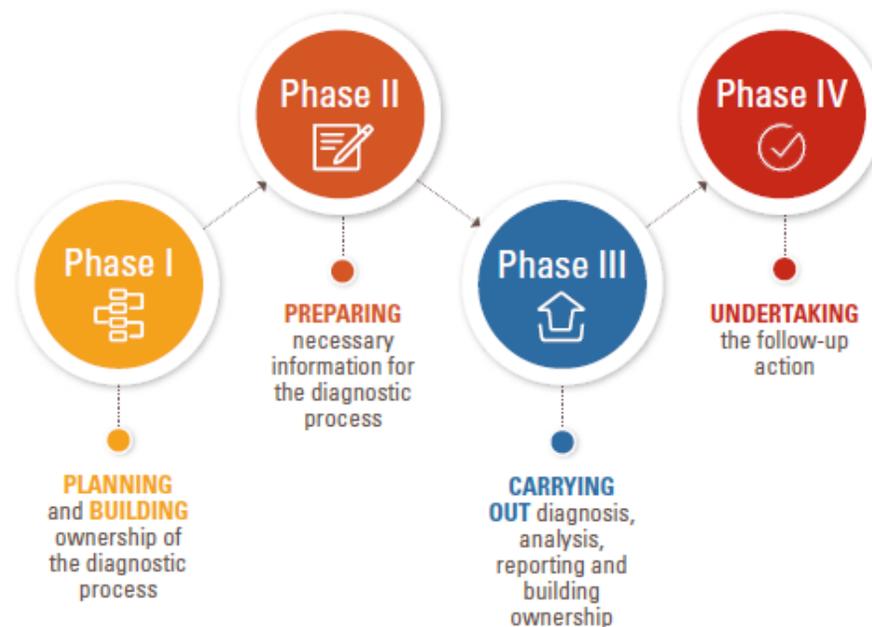
- facilitates learning for all key stakeholders.

- **Focusing on stakeholder utility**: Ultimately, all steps of the diagnostic process have to be oriented towards stakeholder utility. This includes using language, evidence and communication channels appropriate for the target audience; formulating findings that are relevant to decision-making; and ensuring that the process incorporates clear follow-up and review processes.

A sequential outline of steps to be taken for the four phases is outlined below. This process is compatible with methodologies for sector capacity assessment and with various governance structures, and it can be led by a government. The main emphasis is on building ownership of the diagnostic process among key stakeholders. The process for conducting the diagnostics is elaborated below.

Starting assumptions

The process is built on two assumptions:

- A government/public sector entity has decided to launch the evaluation diagnostic process, either because of its own interests and incentives or because it needs to respond to demands and pressures.

- The entity has conducted a readiness assessment for evaluation capacity diagnosis and concluded it is worth proceeding.

Phase I: Planning and building ownership of the evaluation diagnostic process

As shown in the figure above, the first phase of the evaluation diagnostic process has five steps. The objective of Phase I is to engage stakeholders and mobilize their support for the process, and to decide who will carry it out and how.

DIAGNOSTIC PROCESS

STARTING ASSUMPTIONS

Phase I — Planning and building ownership of the diagnostic process

- Engage stakeholders on the need for a diagnostic process
- Establish an oversight team
- Develop a concept note
- Launch the diagnostic process

Phase II — Preparing necessary information for the diagnostic process

- Collect data
- Analyse data

Phase III — Carrying out diagnosis, analysis, reporting and building ownership

- Undertake diagnosis
- Draft report
- Peer review and refine findings
- Publish and disseminate the final report

Phase IV — Undertaking the follow-up action

- Establish a follow-up structure
- Develop and implement an action plan
- Monitor and follow-up

START OVER

1. **Engage stakeholders on the need for a national evaluation capacity diagnostic process and for strengthening national evaluation capacities.** This starting point usually precedes the formal diagnostic process. Since the entities may not be well-resourced or powerful within the government architecture, the primary objective of this dialogue is usually to obtain support for the process from senior staff, such as ministers or the office of the president or head of the entity. This may require 'selling' the evaluation to senior government stakeholders who may not fully recognize its purpose, or may conflate it with monitoring, or may fear or resent it. A secondary objective is to scan stakeholders to determine if there are interested actors elsewhere who will want to be involved and who might use the results. This might involve consultations with other agencies and levels of administration, and with civil society, the private sector and international development partners. An exhaustive mapping is not necessarily required at this stage, as the emphasis remains on generating enough support for starting the process. However, at this stage it is worth considering possible interlinkages and cross-sectoral collaboration.

2. **Establish an oversight team and assign responsibilities for the process** An Oversight Team has overall responsibility for the diagnostic process. The precise structure of its responsibilities varies according to the approach taken. For decentralized evaluation diagnostics, the Oversight Team will likely lead, design and execute the diagnostic process. Where the process takes place at the national level, a broad-based Collaborative Review Group should also be established. This consists of higher level stakeholders who support the process but lack the time to be significantly involved. For more involved processes, the Oversight Team should involve high-level stakeholders in an oversight and governance role, and delegate the actual assessment process to an Assessment Team. It is generally a good practice to have both a technical team (with a team leader) and a wider consultative group of stakeholders, which should ideally involve at least a senior government official supporting the process (an institutional sponsor, or several for a jointly led model). For rapid assessments, a senior official should at least provide a letter of support or introduction to endorse the process and facilitate the team's access to stakeholders.

3. **Develop a draft concept note covering objectives, methods and resources needed.** The concept note sets out key characteristics of the diagnostic process. It might include, for example, a list of key issues to be agreed on such as objectives/purpose, intended audience, scope of the process (unit of analysis), management arrangements and roles of various stakeholders, financing of the assessment, and overview of the intended methodology and information requirements. The draft concept note should be shared with key stakeholders for review, an important step because the concept note defines the rest of the process. The note itself should be brief, rarely exceeding six pages. For quicker diagnostic processes, it may be considerably shorter, or not formally be written at all (even when not written, however, the key questions should be addressed and key stakeholders consulted).

10

Launch the diagnostic process. An in-depth diagnostic process may involve launches across government entities and at different levels, while a decentralized process may be a smaller scale.

Phase II: Preparing necessary information for the diagnostic process

The second phase of the diagnostic process involves preparing a preliminary diagnostic for each national evaluation capacity dimension outlined in the modules. Although this phase contains few steps, it is likely to be one of the most intensive phases, as well as the one exposed to the most important risks, surprises and delays.

1. **Collect data.** This step includes collecting data for responding to the questions in the modules. Where back-up evidence is not readily available the process of completing the modules may be longer. For some areas consultations with other entities and levels of government may be needed. Adequate time should be set aside for consultations and for substantiating answers.

2. **Analyse the data.** Analysing the data requires triangulation of information collected (verifying that information from different sources coincides). For instance, stakeholder perceptions about the number and quality of evaluations can be checked against actual evaluation reports sampled. This analysis should also highlight data gaps requiring further data collection. This step ensures that the diagnosis provides an accurate representation of the needs and challenges facing the programme/entity. (Where time does not allow this step, these will be identified as limitations of the diagnostic process).

Phase III: Carrying out diagnosis, analysis, reporting and building ownership of the diagnostic findings

The third phase of the process consists of four steps: diagnosing evaluation capacities; drafting the diagnostic report; holding a peer review and refining the findings; and publishing and disseminating the final report. The objective of this phase is to provide a summary overview of national evaluation capacities, and to disseminate this information in a useful way to key stakeholders.

1. **Undertake diagnosis of the evaluation capacities.** This also involves choosing actions for an evaluation framework. These activities will be carried out using the four modules in sections 5 through 8.

2. **Prepare the draft report.** Usually the draft report is prepared by the individuals involved in data collection and analysis. It summarizes findings on the four dimensions of national evaluation capacity and summarizes diagnostic and relevant contextual information. The report

should respond to the objectives outlined in the concept note and be written with the target audience in mind. It should also include a section on recommended follow-up action.

Undertake peer review and refine the findings. The draft report should be shared with stakeholders (e.g. the Consultative Review Group, if one was established) to inform them of preliminary findings, give them an opportunity to provide feedback, and ensure a degree of independent quality assessment. If time allows, this peer review can be conducted in several rounds, first by soliciting detailed comments from peers or technical experts and then by disseminating it more widely to higher level reviewers. If resources are available, a one-day workshop is useful for a detailed discussion of the report with key stakeholders

3. **Publish and disseminate the final report.** This takes place after the report is revised based on feedback. If resources are available, a formal launch event can be held. To prepare follow-up action, any event or communication about the final report should outline the upcoming follow-up process and invite stakeholders to participate in it.

Phase IV: Undertaking the follow-up action

The fourth phase of the diagnostic process consists of three steps: engaging stakeholders on the need for follow-up and establishing a follow-up structure; developing an action plan; and monitoring and following up on the action plan. The objective of this phase is to leverage the knowledge and momentum produced by the diagnostic process to build national evaluation processes and systems at the level where the diagnostic process will take place. This phase is closely tied to and likely to overlap with the previous phase, particularly if the diagnostic model chosen faces limited time and resources.

1. **Establish an entity to oversee the follow-up process**. This may be the same team as the one that oversaw the diagnostic process, but rather than taking on a technical role the team can now be thought of as a process enabler and can include an evaluation practice group on national evaluations. Membership can be flexible, but to facilitate the reform process it helps to include institutions affected by intended changes, as well as institutions spearheading or financing the process. In order to avoid diluting responsibility, the follow-up entity should have a structure for designating a team leader or chair who is formally responsible for the follow-up process.

2. **Develop and implement the action plan**: The evaluation practice group should lead the development of an action plan that identifies feasible priorities, clear targets, a timeline and required resources (financial, human and political). It is crucial to build ownership of the action plan by developing it in a consultative manner, to avoid presenting institutions with priorities or reform actions they did not

anticipate or do not support. Consultations can take place through a workshop or several rounds of feedback on a draft, depending on resources available.

3. **Monitor and follow up on the action plan**: Once the action plan is launched, focal points should be identified in key reform units/institutions to report on progress against the targets outlined in the action plan, and to identify challenges encountered and good practices established. Information on progress and issues should be regularly reported via the community of practice to key stakeholders who participated in the diagnostic process. The action plan can be updated as needed over time. Following an adequate period of monitoring and review (usually three to five years), a renewed diagnostic process should be considered to examine progress more comprehensively.

E. How is the Tool implemented?

The Tool has four modules that evolve through a series of steps to assess key evaluation bottlenecks and specific needs and develop context-specific evaluation parameters. The approach is non-prescriptive to account for institutional and other development context differences.

The Self-Assessment Tool begins with a few questions on readiness and preparation for the diagnostic tool, quick assessment (stock-taking) of the policy environment and institutional arrangements. Next a brief profile is developed of the national entity/programme for which the evaluation diagnosis is being carried out. Finally come the four detailed modules.

Module 1: Understanding the enabling environment, which is fundamental to developing national evaluation systems (Hyperlink to the module)

Module 2: Connecting national systems and Agenda 2030 is central to national evaluation systems (Hyperlink to the module)

Module 3: Strengthening and institutionalizing evaluation capacities (Hyperlink to the module)

Module 4: Integrating the SDGs / sustainable development specificities into evaluation processes and evaluations (Hyperlink to the module)

The modules include report cards to be completed that will aid in periodic assessment of progress.
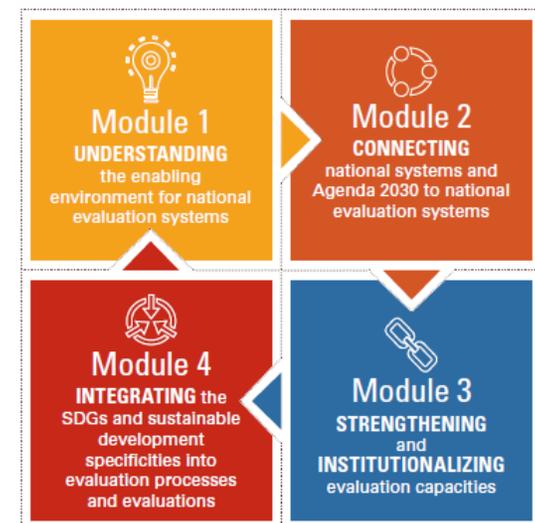
2. PREPARATIONS FOR AN EVALUATION CAPACITY DIAGNOSIS

This section aids the entity in assessing its readiness and level of preparation for undertaking the self-assessment. This will enable the entity to clarify its purpose, determine the resources required, engage stakeholders and ensure access to the data needed to carry out the diagnostic process and develop an evaluation framework. Answering these questions will help the entity to decide whether it needs to spend more time on planning and building ownership of the diagnostic process (Phase 1) or is ready to undertake the assessment.

1. The entity assigned to conduct a readiness assessment: (a) conducted the assessment and concluded that it is worth proceeding; (b) addressed key requirements and concluded that it is worth proceeding; (c) did not conduct a readiness assessment.

2. Clarity on the purpose of the evaluation diagnostic process: (a) has been fully ascertained; (b) has been partly ascertained; (c) has yet to be ascertained.

3. A concept note outlining the objectives, methods and resources needed for evaluation is: (a) fully prepared; (b) partly prepared; (c) yet to be prepared; (d) not needed.

4. Steps to identify key stakeholders who need to be engaged in the diagnostic exercise are: (a) fully taken; (b) partly taken; (c) yet to be ascertained.

5. Steps to engage key stakeholders in the diagnostic exercise are: (a) fully taken; (b) partly taken; (c) yet to be ascertained.

6. An oversight team to lead, design and execute the diagnostic process and assign responsibilities is: (a) fully established; (b) partly established; (c) yet to be established; (d) not needed.

7. Senior-level commitment to the diagnostic process is: (a) fully ensured; (b) partly ensured; (c) yet to be established.

8. An institutional review process is: (a) fully established; (b) partly established; (c) yet to be established; (d) not needed.

9. The back-up data for responding to the questions in the diagnosis modules are: (a) fully sufficient; (b) partly sufficient; (c) insufficient.

10. The instruments for wider stakeholder consultations to ascertain evaluation status and need are: (a) completely finalized; (b) partly finalized; (c) yet to be established; (d) not needed.

11. A review process to provide feedback on the diagnosis and next steps for strengthening evaluation processes and systems is: (a) fully established; (b) partly established; (c) yet to be established.

3. A QUICK ASSESSMENT OF THE POLICY ENVIRONMENT

A. Policies and institutional arrangements for development data and statistics

1. Policies for comprehensive development data and national statistics are: (a) fully in place; (b) partly in place; (c) not yet in place.

2. Institutional arrangements for development data and national statistics are: (a) fully in place; (b) partly in place; (c) not yet in place.

If institutional arrangements are fully in place:

3. In terms of the quality of the institutional arrangements for data and statistics, they are: (a) robust and fully functional; (b) partly robust and partly functional; (c) not robust and not yet functional.

4. Knowledge of statistical standards (for example, concepts and definitions) and implementation of them are: (a) not an issue; (b) partly an issue; (c) a major issue.

B. National monitoring systems

1. Monitoring systems for ascertaining progress on the SDGs are: (a) fully established; (b) partly established; (c) yet to be established.

> Any further elaboration on the above areas:

2. Monitoring systems to track progress on development results are: (a) fully functional; (b) partly functional; (c) yet to be constituted.

15

If monitoring systems are fully/partly in place:

3. In terms of quality, monitoring systems are: (a) robust and fully functional; (b) partly robust and partly functional; (c) not robust and not yet functional.

4. Institutional arrangements for reporting on the SDGs are: (a) fully established; (b) partly established; (c) not yet established.

5. In economic and social development policy development, statistics are: (a) widely used; (b) partly used; (c) not used.

6. Allocation of resources for statistical activities is: (a) not an issue; (b) partly an issue; (c) a major issue.

7. Availability of baseline data on a majority of SDG indicators and disaggregated data is: (a) not an issue; (b) partly an issue; (c) a major issue.

Any additional information on availability and quality of data and statistics:

8. Design and implementation of appropriate methodologies and techniques for measuring development indicators are: (a) not an issue; (b) partly an issue; (c) a major issue.

9. Coordination between statistical agencies and other development data producers is (a) not an issue; (b) partly an issue; (c) a major issue.

10. Administrative data systems are (a) well developed; (b) partly developed; (c) not yet developed.

Periodic collection of development data is: (a) not an issue; (b) partly an issue; (c) a major issue.

C. Institutional arrangements for making and implementing policies

Functions of central government entities

Central government entities:

1. Shape high-level policy (understanding issues and possible solutions): (a) Completely shape; (b) partly shape; (c) do not shape.

2. Track the performance of national policy priorities: (a) Fully track; (b) partly track; (c) do not track.

3. Track the performance of reforms and new/innovative policies and programmes: (a) Fully track; (b) partly track; (c) do not track.

4. Ensure consistent implementation of national mandates at subnational and local levels: (a) Fully ensure; (b) partly ensure; (c) do not ensure.

5. Manage quality assurance and harmonization: (a) Fully manage; (b) partly manage; (c.) do not manage.

6. Identify development programme inefficiencies: (a) Fully identify; (b) partly identify; (c) do not identify.

7. Identify development spending inefficiencies: (a) Fully identify; (b) partly identify; (c) do not indentify.

8. Demonstrate performance in developing and implementing policies: (a) Fully demonstrate; (b) partly demonstrate; (c) do not demonstrate.

9. Further a culture of accountability and transparency: (a) Fully further; (b) partly further; (c) do not further.

For countries with federal systems: Continue

There are state or subnational government entities that fulfil the following functions:

10. Shape state-level policy (understanding issues and possible solutions): (a) Completely shape; (b) partly shape; (c) do not shape.

11. Establish vertical linkages with federal entities and policies: (a) Fully establish; (b) partly establish; (c) do not establish.

12. Track the performance of state/subnational policy priorities: (a) Fully track; (b) partly track; (c) do not track.

13. Track the performance of reforms and new/innovative policies and programmes: (a) Fully track; (b) partly track; (c) do not track.

14. Ensure consistent implementation of national mandates at subnational and local levels: (a) Fully ensure; (b) partly ensure; (c) do not ensure.

15. Manage quality assurance and harmonization: (a) Fully manage; (b) partly manage; (c) do not manage.

16. Identify inefficiencies: (a) Fully identify; (b) partly identify; (c) do not identify.

17

17. Demonstrate performance in making and implementing development-related policies: (a) Fully demonstrate; (b) partly demonstrate; (c) do not demonstrate.

Further a culture of accountability and transparency: (a) Fully further; (b) partly further; (c) do not further.

D. Contextual factors affecting evaluation

Contextual factors have a bearing on evaluation systems and processes and conduct of evaluations. Evaluations require planning, expertise, resources and time. They are thus less likely to get the attention they need in rapidly evolving environments that require quick decision-making; they are more adapted to informing stable and long-term initiatives. On the other hand, evaluation plays a legitimizing function for governments, since it either demonstrates positive performance or at a minimum the government's willingness to learn from mistakes and improve services for citizens. Crises can thus create demand for performance-oriented feedback mechanisms such as evaluation. The larger and more complex the affairs administered by the public sector, the more important the role of evaluation to guide them.

1. Are there any contextual factors that would impede evaluation efforts? If yes, what are those factors?

   a) Political conflict.
   b) Conflict with other countries.
   c) Other factors.

2. If there are contextual factors that impede evaluation efforts, to what extent do these factors impact development processes?

   a) Significantly.

   b) Moderately.

   c) Insignificantly.

3. If yes to question 2, to what extent do they impact the evaluations that will be undertaken?

   a) Significantly.

   b) Moderately.

18

c) Insignificantly.

Summary of context:

    A. Ensure there is adequate supporting data for the responses provided. (Insert summaries of key data here).
    B. Cross-check with relevant actors and provide their perspectives, both supporting and contrary.
    C. Provide a brief note on key actors consulted in assessing institutional context.

Certain evaluation methods are specifically designed for dynamic or humanitarian contexts
(see http://www.betterevaluation.org/en/plan/approach/developmental_evaluation) and real-time evaluation (see
http://www.alnap.org/what-we-do/evaluation/rte).

4. PROFILE OF THE NATIONAL ENTITY/PROGRAMME UNDERGOING EVALUATION DIAGNOSIS

A. Institutional profile

1. Name of the entity/programme seeking to diagnose evaluation capacities.

2. Level of government entity/national programme that is seeking to assess its evaluation capacities.

   National evaluations entity:

   a) National development planning entity
   b) Ministry (drop-down tool for ministry details)
   c) Specific sector (drop-down tool for sector details)

d) Local government

e) Specific programme or project (drop-down tool for details)

f) Others (details)

3. What is the mandate of the entity/national programme seeking to assess evaluation capacities?

> A brief description of the mandate of the entity/national programme, and its positioning compared to other institutions.

B. Expectations from the diagnosis

1. What are the expectations in diagnosing evaluation capacities? (Select responses that reflect your expectations)

    a) Assess current bottlenecks and needs and develop an evaluation strategy for the entity/programme.

    b) Assess current bottlenecks and needs and develop an evaluation strategy for the national development plan.

    c) Constitute an evaluation entity and centralized evaluation function.

    d) Mainstream evaluation across all levels and departments of government, and beyond.

    e) Simply evaluate key evaluation bottlenecks and needs.

    f) Create a culture of consultation and critical review of performance.

    g) Others (specify).

2. Is there commitment at the senior level to the diagnostic process and to developing an evaluation framework? (a) Full commitment; (b) partial commitment; (c) no commitment yet.

3. Is there a commitment at the senior level in your entity to prioritizing evaluation and committing resources? (a) Full commitment; (b) partial commitment; (c) no commitment yet.

MODULES

The Tool has four modules to help unpack requirements for building evaluation capacity. The approach is non-prescriptive to account for different institutional and development contexts. Depending on the needs of the entity/programme, it may use all four modules, or select one or more that are the most relevant.

4. MODULE 1: UNDERSTANDING THE ENABLING ENVIRONMENT

A. Need for evaluation

The extent of need for systematic evaluations and sound evaluation practices varies from country to country. Evaluations of government strategies and programmes have benefits, such as informing and legitimizing decisions, but they also bear costs, such as time, money and effort. Government entities should not hesitate to invest resources in formal evaluative processes and systems if current arrangements do not satisfy their needs for information and legitimacy. However, if the current evaluation arrangements satisfy the needs for information and legitimacy, the national entity should determine whether or not they wish to invest the time and human and financial resources required to strengthen and formalize the national evaluation function.

Assess the need for evaluative information and the added value of evaluative practices by answering the following questions. Answers to the questions do not linearly indicate a greater or lesser need for evaluative practices, but convey different implications about what formalized evaluation may or may not contribute. The questions also help to assess whether the evaluation is relevant to the issues and decisions at hand.

1. Do you know what is working and what is not in your entity/programme? Are you: (a) fully informed; (b) moderately informed; (c) not informed.

2. Are there political considerations that need to be accounted for? Examples include: development areas that may be politically sensitive, leading to greater resistance to evaluation; power equations that may undermine conduct of evaluations; country context factors such as conflict in which the solution is beyond the scope of your entity; or resource issues. (a) There are no political considerations; (b) certain political considerations exist; (c) many political consideration need to be accounted for.

3. Does your entity/national programme entail complex areas needing multisector and multi-agency engagement? (a) There are no complex areas; (b) some complex areas need multisector and multi-agency engagement; (c) most areas need multisector and multi-agency engagement.

4. In your entity/programme, is there a clear sense of what defines success and failure? There is: (a) a clear sense of what defines success and failure; (b) a moderate sense of what defines success and failure; (c) no understanding of what defines success and failure.

5. How frequently do you feel you need information to make good decisions or track the progress of ongoing efforts? (a) Frequently need information; (b) need information on an ad hoc basis; (c) have limited need for information.

6. Do existing feedback mechanisms fulfil the need for information to make good decisions or track the progress of ongoing efforts? (a) Completely fulfil; (b) partially fulfil; (c) do not fulfil.

7. To what extent do key stakeholders of your entity/programme agree on what is working and what needs to change? (a) Are mostly in consensus; (b) are partially in consensus; (c) are not in consensus.

8. To what extent do stakeholders of your entity/programme trust each other to work towards the same objectives? (a) Mostly trust each other; (b) partially trust each other; (c) do not trust each other.

9. To what extent do you know what other government entities think about the performance of your entity/programme? (a) Know to a large extent; (b) know to some extent; (c) are not aware.

10. To what extent do you know what external stakeholders think about your entity/programme performance? (a) Know to a large extent; (b) know to some extent; (c) are not aware.

11. Do external stakeholders demand justification for your actions or your existence? (a) To a large extent; (b) to some extent; (c) not at all.

12. Is your entity/programme recognized as a key government entity? (a) To a large extent; (b) to some extent; (c) not recognized as a key entity.

Further information pertaining to need for evaluations:

B. Identifying evaluation needs of key stakeholders

Many government actors at different levels are involved in national development efforts, and they have with diverse roles and responsibilities. Evaluative practices, therefore, should provide them with the information needed to fulfil diverse purposes and respond to different organizational interests. It is helpful to disaggregate the government and the public sector into their constituent entities to identify how their interests relate to different levels of evaluative practice. Linking stakeholder relations to evaluations helps to clarify what interests or opportunities evaluation serves for different stakeholder groups, and what kind of risk, threat or cost it carries for them.

Identify the key utility of evaluation for major actors. Add or remove stakeholder groups to reflect the national context.

1. The extent to which the parliament is interested in understanding whether approved policies have the intended effect: (a) Fully interested; (b) partially interested; (c) not yet interested.

2. The extent to which the parliament is interested in more evaluations of development outcomes: (a) Fully interested; (b) partially interested; (c) not yet interested.

3. The extent to which political parties are interested in understanding whether approved policies have the intended effect: (a) Fully interested; (b) partially interested; (c) not yet interested.

4. The extent to which citizens are interested in understanding whether approved policies have the intended effect: (a) Fully interested; (b) partially interested; (c) yet to be interested.

5. The extent to which national civil society organizations/think tanks are interested in understanding whether approved policies have the intended effect: (a) Fully interested; (b) partially interested; (c) yet to be interested.

6. The extent to which the international development community is interested in understanding whether approved policies have the intended effect: (a) Fully interested; (b) partially interested; (c) yet to be interested.

Pop-up matrix:

Identifying stakeholder interests: What makes evaluation useful and relevant to different national stakeholders?

Stakeholder group	Uses of evaluation
Central government (executive branch)	• Shape high-level policy (understanding issues and possible solutions) • Track the performance of national policy priorities (e.g. poverty reduction, such as Uganda in Rist 2004) • Track the performance of reforms and innovative policies and programmes • Undertake control, quality assurance and harmonization; ensure consistent implementation of national mandates at subnational and local levels • Identify inefficiencies • Demonstrate good performance • Further a culture of accountability and transparency (e.g. Malaysia in Rist 2004)
Parliament (legislative branch)	• Accountability: Determine whether approved policies are having the intended effect • Fulfil desire for more timely evaluations of development outcomes
Institutional and middle management (public servants)	• Implementation: Understand how organization, management, resource allocation, implementation, operations could be improved
Front-line service-providers, street-level bureaucrats, practitioners	• Customer satisfaction: Ensure that front-line realities and citizen satisfaction are fed back into decision-making • Fulfil desire for more frequent, formative evaluations of processes
Political parties, the media, citizens at large	• Obtain detailed information about government performance to make informed decisions about renewing political mandates

C. Demand for evaluation

1. Regarding demand for evaluation which is/are a priority (tick one or more)

   a) More frequent evaluations to adjust processes.
   b) Evaluations of progress on development outcomes.
   c) Evaluations of citizens' satisfaction.
   d) Obtaining detailed information about government performance to make informed decisions about renewing political mandates.

24

e) Accountability to key stakeholders.

f) Others.

2. A majority of the key stakeholders of the entity/programme consider evaluations useful for improving performance: (a) Agree; (b) partly agree; (c) disagree.

3. A majority of key stakeholders of the entity/programme view evaluations as useful for improving accountability: (a) Agree; (b) partly agree; (c) disagree.

4. A majority of key stakeholders view evaluations as resource intensive and lack comparable utility: (a) Agree; (b) partly agree; (c) disagree.

5. A majority of key stakeholders fully agree that evaluations are useful in the country context: (a) Agree; (b) partly agree; (c) disagree.

Summary of stakeholders and demand for evaluations:

D. National evaluation policies and institutional arrangements

1. National evaluation policy: (a) There is a national evaluation policy; (b) although there is no evaluation policy, there are government guidelines for evaluation; (c) there is no evaluation policy or formal guidelines.

Where there is a national evaluation policy:

2. The policy is: (a) fully implemented; (b) partly implemented; (c) not implemented.

3. Institutional arrangements for national evaluations are: (a) fully in place; (b) partly in place; (c) not in place.

4. The evaluation function: (a) is located in a centralized stand-alone evaluation unit; (b) is located in a multifunctional division with a dedicated evaluation unit; (c) is not supported by institutional arrangements.

5. Institutional arrangements for evaluations in sector ministries: (a) Policies and institutional arrangements are in place; (b) some sector ministries have policies and institutional arrangements in place; (c) policies and institutional arrangements are not in place.

6. Systematic prioritization of development evaluations: (a) There is systematic prioritization; (b) there is moderate prioritization; (c) there is no prioritization.

7. Financial resources assigned for evaluations: (a) Adequate financial resources; (b) modest financial resources; (c) no financial resources.

8. Human resources assigned for evaluations: (a) Adequate human resources; (b) modest human resources; (c) no human resources.

9. Adequacy of management and implementation arrangements in place for managing evaluations: (a) Fully adequate;
    (b) partly adequate; (c) not adequate.

10. Adequacy of human resources assigned for managing and implementing evaluations: (a) Fully adequate; (b) partly adequate; (c) not adequate.

11. Adequacy of arrangements in place for conduct of evaluations: (a) Fully adequate; (b) partly adequate; (c) not adequate.

12. Adequacy of funding available for evaluations: (a) Fully adequate; (b) partly adequate; (c) not adequate.

Summary of institutional context

Overall score of the level of evaluation capacities based on the assessment/stock-taking of data systems, monitoring and evaluation.

- Mature
- Evolving
- Emergent

Recheck to see if you agree with the overall score and categorization of the level of national evaluation.

E. Accelerators of support for national evaluation

Contextual factors that create the enabling environment (e.g., resources, political actors, support for change, etc.) are usually relatively fixed in the short run. Nonetheless, lessons from countries show that evaluation practices have grown and evolved, sometimes rapidly. Recognizing that countries' experiences are diverse and there is more than one path to developing national evaluation capacities is an important lesson from decades of development evaluation experience.

Evaluation systems and processes develop and consolidate through different pathways. The evolution of evaluation serves as a map of possibilities and entry points, and it can help identify experiences that might resemble or inspire your situation.

1. Which of the following best describes your country/entity context in terms of support for the growth of evaluation? Choose all responses that apply.

   a) Evaluation is valued although there is no proactive emphasis on it (positive consideration).
   b) There is demand for performance information from the top. For example, higher level government officials/members of parliament seek evaluation information to test impact, demonstrate performance or implement a results-oriented agenda (demand from the top).
   c) Ongoing systemic change and reform provide opportunities to integrate evaluation principles and practices into new structures (windows of opportunity).
   d) There is demand from the community for accountability and performance (citizen demand).
   e) Evaluations taking place in individual sectors or programmes are successful, setting a precedent for gradual scale-up of evaluation practice (lessons from other areas of the government).

F. Engaging with partners with comparable objectives

The development of national evaluation capacities is a long-term endeavour. Evaluation capacity is closely tied to national public-sector capacity, and cannot exceed it. Evaluation capacities are built organically, unevenly and over time, and countries with governance challenges are less likely to leverage evaluative information. Long-term evaluation capacity development rests on the ability to craft long-term solutions, which are more effective than short-term capacity gains or the use of foreign models. Also, to consolidate evaluation

27

capacities, it is critical to develop strong links with national institutions that have an accountability mandate, for example, audit units and statistical and monitoring agencies.

1. How likely are the following evaluation partnerships?

   a) With the national audit entity: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   b) With national statistical systems: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable

   c) With performance monitoring entities: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   d) With the parliament: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   e) With civil society organizations/think tanks/NGOs advocating for good public services, efficient use of public funds, and transparency and access to information: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   f) With civil society organizations advocating for evaluation, e.g., voluntary organizations for professional evaluators (VOPEs): (a) Likely; (b) moderately likely; (c) not likely; (d) not applicable.

   g) With inter-country peer-to-peer evaluation collaborators (for knowledge sharing and support): (a)Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   h) With regional evaluation forums: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   i) With global evaluation forums: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   j) With donor agencies: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

   k) With multilateral agencies and their platforms: (a) Very likely; (b) moderately likely; (c) not likely; (d) not applicable.

2. Which partnerships need further consolidation for your entity/national programme to build or strengthen evaluation capacities?

a) With the national audit entity.
b) With national statistical systems.
c) With performance monitoring entities.
d) With parliament.
e) With civil society organizations, think tanks and NGOs advocating for good public services, efficient use of public funds, and transparency and access to information.
f) With civil society organizations advocating for evaluation (e.g., VOPEs).
g) With inter-country peer-to-peer evaluation collaborators (for knowledge sharing and support).
h) With regional evaluation forums.
i) With global evaluation forums.
j) With donor agencies.
k) With multilateral agencies and their platforms.

If the entity wishes to carry out a survey to assess the national evaluation capacities across government entities click here: Link to the Survey

Summary of partnerships:

A pop-up window of examples:

- The (hyperlink National Council for the Evaluation of Social Development Policy (CONEVAL)) in Mexico is an autonomous constitutional evaluation entity. CONEVAL evaluates the impact of priority public policies on poverty reduction through an integrated, multidimensional poverty index that captures progress in education, health care, social security, etc. This strong and independent evaluation function was largely built on the results of a select number of social policy impact evaluations in the 1990s. These evaluations involved randomized controlled trials of a World Bank-sponsored social safety net programme, Progresa, that provides conditional cash transfers to poor households. The evaluations successfully demonstrated the effectiveness of the programme and legitimized its continuation even following a change in government. This demonstrated the value of evaluation as an instrument for effective policy making and launched the successful CONEVAL initiative.

- Bottom-up pressure: Pressure from citizen groups and civil society organizations for evaluation has supported citizen participation in accountability. Examples include evaluation forums such as Africa Evaluation Association (AfrEA) and Global Parliamentarians Forum for Evaluation.

- Peer collaboration and South-South collaboration: Exchanges on relevant experiences and lessons among countries help to develop national evaluation systems, for example between Uganda and Benin in the Twende Mbele Programme (see below).

- The pop-up window for Twende Mbele: Twende Mbele Programme is a South-South collaboration to enhance monitoring and evaluation systems in Uganda, South Africa and Benin. The three countries have agreed to share experiences to improve their monitoring and evaluation (M&E) systems and implementation practices, policies, tools and procedures. The programme consists of active collaboration, not just networking.

  Examples of collaborations currently being undertaken are: (a) Adapting the Management Performance Assessment Tool as used in South Africa and planning for taking it forward in Uganda and Benin; (b) a workshop involving the three countries to reflect on opportunities to involve civil society in strengthening the national M&E systems; (c) a diagnosis of the gender responsiveness of the countries' national M&E systems and how it can be strengthened; and (d) a review of evaluation curricula of university courses across Africa and the potential to collaborate.

- Piggy-backing on other systems: Evaluative practices are not developing as an independent function but under the helm of established feedback mechanisms, which provide a sheltered incubator and institutional legitimacy, for example in Brazil (performance audits) and Nigeria (national monitoring).

Summary of the enabling environment:

A. Ensure there is adequate supporting data for the responses provided, and summarize key data here.

B. Cross-check with relevant actors and provide their perspectives, both supporting and contrary.

C. Provide a brief note on key actors consulted in assessing institutional context.

6. MODULE 2: CONNECTING NATIONAL SYSTEMS AND AGENDA 2030 TO NATIONAL EVALUATION SYSTEMS

This module facilitates taking stock of links between national systems and the SDGs. An understanding of these links is critical for framing national evaluations.

'National systems' as used here implies a web of values, institutions, incentives, policies, plans and priorities that determine how decision-making, planning, implementation and accountability processes work in a country. Most countries operate through a multiplicity of systems and authorities, such as municipal, provincial, state and sectoral levels of governance.

This module does not provide guidance for a comprehensive approach to integrating the SDGs at a national level; other guides have been produced for this purpose, such as the United Nations Sustainable Development Solutions Network's SDG Guide (*Getting Started with the Sustainable Development Goals. A Guide for Stakeholders*, 2015) and MAPS (https://undg.org/document/maps-mainstreaming-acceleration-and-policy-support-for-the-2030-agenda/).

Key steps:
   A. Promoting Agenda 2030 nationally (Hyperlink to the section).
   B. Mapping national plans against the SDGs (Hyperlink to the section).
   C. Integrating the SDGs into national and sectoral plans (Hyperlink to the section).
   D. Assigning coordination responsibility (Hyperlink to the section).
   E. Mainstreaming the SDGs into entities at national and subnational levels
   F. (Hyperlink to the section)

A. Promoting Agenda 2030 nationally

If countries are to invest in follow-up and review processes for SDG targets and principles, they must first create awareness of them among citizens.

1. Has the government taken measures to increase awareness about the SDGs among citizens? Such measures have been: (a) fully taken; (b) partly taken; (c) not taken.
2. Are there forums and mechanisms to enable discussions on prioritization of the SDGs and targets? (a) Multiple forums and mechanisms exist; (b) few forums and mechanisms exist; (c) no forums or mechanisms exist yet.
3. Has the entity/national programme taken measures to increase awareness about the SDGs among citizens? Such measures: (a) have been fully taken; (b) have been partly taken; (c) are yet to be taken.

B. Mapping national plans against the SDGs

When the SDGs were adopted in 2015, many countries were in the process of implementing various national and sectoral strategies, ranging from short term to long term. In order to acknowledge the 2030 agenda without disrupting ongoing national implementation of such strategies, many governments have chosen to map their existing national and sectoral development objectives and indicators against the 17 SDGs. This enables the identification of gaps, discrepancies and areas of convergence.

1. Has the government taken measures to map national plans against SDG priorities? Such measures have been: (a) fully taken; (b) partly taken; (c) not yet taken.

C. Integrating the SDGs into national and sectoral plans

1. Has the government prioritized integrating the SDGs into national plans and strategies? The government has: (a) fully prioritized integration; (b) partly prioritized integration; (c) not yet prioritized integration.

2. Has the government taken measures to integrate the SDGs into national and sectoral plans? The government has: (a) fully taken such measures; (b) partly taken such measures; (c) not yet taken such measures.

3. Has the government taken measures to adopt a more integrated approach to development planning and implementation? The government: (a) has fully taken such measures; (b) has partly taken such measures; (c) has not yet taken such measures.

Pop-up reference: Guidance: Chapter 3 of Sustainable Development Solutions Networks' SDG Guide, 'Tools for designing SDG strategies and roadmaps', includes references to needs assessment and costing for individual goals.

Pop-up example: How Kenya is connecting its national systems to the SDGs. Kenya has prepared a roadmap that maps the 17 SDGs with its Vision 2030 Second Medium Term Plan objectives. Kenya also plans to integrate the SDGs into its third Medium Term Plan, covering 2018-2022. The government has directed all ministries, departments and agencies to mainstream the SDGs into policy, planning, budgeting, and monitoring and evaluation systems and processes, and to ensure that planning and budgeting officials have been trained to do so.

At the institutional level, SDG implementation and monitoring are coordinated by the Ministry of Devolution and Planning. The SDGs focal point in the Ministry offers technical backstopping for SDGs within the government and among external stakeholders. An inter-agency technical committee has been set up, with members from line ministries, the National Bureau of Statistics, the National Council for Population and Development, civil society organizations and the private sector.

On monitoring and reporting, the Kenya National Bureau of Statistics has undertaken indicator mapping, identifying 128 indicators out of the 230 global SDG indicators that can be measured with available data or data that can be produced within one to two years with minimum effort. (Source: Risse, Nathalie (2017), VNR Main Messages Highlight Diverse SDG Approaches, International Institute for Sustainable Development (IISD), http://sdg.iisd.org/news/vnr-main-messages-highlight-diverse-sdg-approaches)

Pop-up on reference: UNDESA's report 'Integrated Approaches to Sustainable Development Planning and Implementation' (2015) provides examples of how Germany, Belize and Uganda have integrated the SDGs into national development strategies.

D. Assigning coordination responsibility

An integrated approach to achievement of the SDGs means that siloed, sector-based approaches to development solutions need to change. Much like the delineation of Agenda 2030 as 17 sectoral or thematic goals, countries operate through a series of entities (ministries, departments, agencies) with sectoral or thematic responsibilities. However, the SDGs have interlinking targets that require a more holistic approach to development. To bridge this gap between a holistic approach and a traditional sectoral approach, some countries have established a high-level entity to lead and coordinate SDG implementation, follow-up and review. These entities are usually attached to central ministries of planning or finance or central political offices such as the prime minister's or president's office. In some countries, these entities are framed as sustainable development councils that oversee national sustainability planning.

1. Has the government established systems to coordinate SDG implementation, follow-up and review? Such systems are: (a) fully established; (b) partly established; (c) yet to be established.

2. Are institutional arrangements in place for coordination of national development planning? (a) Institutional arrangements are in place; (b) Institutional arrangements are in place for relevant areas; (c) No institutional arrangements have been identified.

Example pop-up note: In Indonesia, the Ministry of National Development Planning (*Bappenas*) is assigned to coordinate the process of planning, implementing, monitoring, evaluating and reporting on the SDGs. The government has appointed a national team to coordinate the SDGs and seeks to involve various stakeholders in it.

Pop-up note on Guidance: SDplanNet's paper on sustainable development councils outlines the role such centralized bodies can play in supporting design and delivery of the SDGs.

E. Mainstreaming the SDGs into entities at national and subnational levels

Mainstreaming the SDGs means updating mandates, plans, policies, procedures, incentives, job descriptions, etc. in line with the Agenda 2030 vision, results and principles of follow-up and review. This is a long-term effort that involves gradual capacity development, as agencies need to be resourced and capacitated to behave differently. Key Agenda 2030 behaviours to be mainstreamed include (a) long-term planning, (b) adaptive management, (c) monitoring and evaluation, and (d) gender and vulnerability analysis.

1. Has the government taken measures to mainstream the SDGs and their principles into policies, strategies, procedures and incentive structures?
   a) Government policies: Such measures are: (a) fully taken; (b) partly taken; (c) not yet taken.

   b) Ministerial strategies: Such measures are: (a) fully taken; (b) partly taken; (c) not yet taken.

   c) Sectoral strategies: Such measures are: (a) fully taken; (b) partly taken; (c) not yet taken.

   d) Government procedures: Such procedures are: (a) fully taken; (b) partly taken; (c) not yet taken.

   e) Government incentive structures: Such measures are: (a) fully taken; (b) partly taken; (c) not yet taken.

2. Has the government undertaken measures for adaptive management, supporting gradual adoption of the SDG approach to development? Such measures are: (a) fully undertaken; (b) partly undertaken; (c) not yet undertaken.

3. Has the government undertaken measures for gender and vulnerability analysis? Such measures are: (a) fully undertaken; (b) partly undertaken; (c) not yet taken.

Summary of the alignment between national development planning and the SDGs, areas where links need to be established as a priority

F. Report card on links between national systems and the SDGs

The report card reflects the overall score in each section, which can be compared with the diagnostic score once the changes have been made.

Report card on links between national systems and the SDGs		
Parameter	**Score: 3 is high, 1 is low, 0 is no activity**	
	Before diagnostic process	**After changes are made**
A. **Promoting Agenda 2030**		
B. **Mapping national plans against SDGs**		
C. **Integrating the SDGs into national and sectoral plans**		
D. **Assigning coordination responsibility**		
E. **Mainstreaming the SDGs into entities at national and subnational levels**		
Total score	{total score value}	{total score value}
Changes in the score		[Date2 – Date 1]

Changes in scores for the selected period will indicate progress as follows. (It is likely that some parameters will not be comparable, and in such cases overall scores will be automatically adjusted.)

Advanced: Mechanisms and instruments for linking national systems and the SDGs are present at an advanced level. Score: >80%
Improved: Mechanisms and instruments for linking national systems and the SDGs are present, sufficient and satisfactory. Score: 70%-<80%

Intermediate: Mechanisms and instruments for linking national systems and the SDGs are present and partially sufficient. Score: 60%-<70%

Incipient: Mechanisms and instruments for linking national systems and the SDGs are moderately present and insufficient. Score: 50%-<60%

Not structured: Mechanisms and instruments for linking national systems and the SDGs are not present. Score: <50
Percentage improvement in the score

Good: Over 50% increase
Medium: 30-50% increase
Poor: Below 30% increase

7. MODULE 3: STRENGTHENING AND INSTITUTIONALIZING EVALUATION CAPACITIES

This module provides instruments and options for strengthening and institutionalizing evaluation capacities. It includes tools designed for integrating evaluation into national and subnational planning and implementation cycles. It also presents key steps in developing a national evaluation system.

Making evaluation ready for the SDGs requires strengthening evaluation capacities at national and subnational levels. Drawing on research and examples from various countries, this module outlines practical steps for integrating M&E practices into national systems.

Two points should be noted. First, these steps are not prescriptive; they illustrate examples of approaches taken, but there are many ways to promote evaluation, and functional M&E systems can come in many different forms. Second, these steps are a means to an end. High-quality evaluations are not valuable in and of themselves. They are just a step on the path to promoting learning, consensus, evidence-driven policymaking and a culture of good governance – and ultimately, sustainable development.

This section addresses both centralized and decentralized systems and processes. The activities differ based on the level of dispersion of decision-making in evaluation for the country or the entity/programme undertaking the diagnostic assessment.

Centralized evaluation system and processes (national government)

A. Assessing the policy environment.

B. Strengthening evaluation policy and function.

C. Mainstreaming evaluation at the ministerial, sector and subnational level

D. Promoting core development values in evaluations.

E. Developing a national evaluation community of practice.

F. Strengthening data and statistics.

G. Conducting periodic evaluation capacity diagnostics.

Decentralized evaluation system and processes (subnational governments: ministries and sectors at the state and local level)

A. Assessing the policy environment.

B. Stregnthening the evaluation framework

C. Mainstreaming evaluations at the ministerial, sector, state and local level

D. Promoting core development values in evaluations.

E. Developing an evaluation community of practice

F. Strengthening data and statistics

Centralized evaluation system and processes

Skip to <u>decentralized evaluation (hyperlink)</u> if the diagnosis is for a ministry, department, sector or programme rather than a national evaluation framework.

A. Assessing the policy environment

This section aims to determine whether the policy environment is right for carrying out a diagnostic process and developing evaluation policy and practices.

1. Will there be legal issues in developing evaluation policy and function? (a) There will not be any legal issues; (b) there may be legal issues, but they can be addressed; (c) legal issues remain a constraint and will take some time to be addressed.

2. Will there be political resistance to more structured evaluation practices? (a) There will not be any political resistance; (b) there may be political resistance, but it can be addressed; (c) political resistance remains a constraint and will take some time to be addressed.

3. Is there political space to use evaluation evidence to bring about positive change? (a) There is political space to use evaluation evidence for change; (b) there is partial political space to use evaluation evidence for change; (c) there is no political space to use evaluation evidence for change.

4. Are there resource constraints in developing evaluation policy and function? (a) There are not any resource constraints; (b) there are some resource constraints, but they can be addressed; (c) resource constraints remain and will take some time to be addressed.

5. Is there technical capacity to develop evaluation policy and function? Technical capacity is: (a) not an issue; (b) partly an issue; (c) a major issue.

B. Strengthening evaluation policy and function

Following are a set of parameters for developing an evaluation policy and functions. The options you choose will be the basis for assessing progress made on each of the parameters.

1. What will the national evaluation policy provide a framework for?
   a) All national entities and programmes.
   b) Primarily the national planning entity.
   c) Selected government entities/ministries and programmes.

If option 'C' is selected, specify the ministries/entities and programmes that will be covered by the evaluation policy.

Pop-up text: Checklist when preparing an evaluation policy.
- Clarify purpose (accountability and learning or accountability alone).
- Outline the concept and role of evaluation.
- Establish evaluation and development planning links.
- Provide an institutional framework.
- Establish reporting lines.
- Outline human and financial resources.
- Establish accountability and learning processes.

Pop-up text for national evaluation policy: Having a national evaluation policy is a key step, although not an essential one, in building evaluation capacities. An explicit evaluation policy should provide a clear explanation of the concept, role and use of evaluation within the organization, including the institutional framework and definition of roles and responsibilities; an explanation of how the evaluation function and evaluations are planned, managed and budgeted; and a clear statement on disclosure and dissemination.

An evaluation policy is usually a written document, formally endorsed by relevant decision-makers, that describes rules and responsibilities for evaluation. An evaluation policy stipulates what evaluation is and is for; what should be evaluated; who should perform the evaluation; how evaluations should be conducted; how resources are assigned for evaluation; etc. National evaluation policies apply at the national level to the entire government. The adoption of an evaluation policy presents certain advantages for the development of evaluation systems, as it clarifies expectations around evaluation; it is public and can be used to hold actors to account or to defend evaluation against political abuse; and it signals that evaluation is approved by higher levels of authority. It also legitimizes development of an evaluation culture.

Pop-up window on examples of national evaluation policies

When Morocco revised its constitution in 2011 it included provisions for evaluation of government initiatives in the revised document. The conduct and use of evaluations are stipulated in various articles (though not in detail, as befits a constitution).

2. What entity will lead and coordinate the evaluation function?

   a) A dedicated and detached (independent) evaluation unit.
   b) A multifunctional division with a dedicated evaluation unit.
   c) There is currently no evaluation unit or coordination entity.

If the response is 'C' skip to question 5

Pop-up text on establishing a dedicated central evaluation unit: A central evaluation unit has the mandate to coordinate and oversee evaluation across all areas of government. Its mandate also includes (a) ensuring that decentralized units fulfil their evaluation duties, (b) gathering data and reports from decentralized M&E activities (c) preparing regular synthesis reports for the executive or legislature, (d) conducting its own national-level evaluations, and (e) contributing to the development of decentralized evaluation capacity. Establishing a central evaluation unit clarifies the national evaluation architecture and creates an institution that parallels national statistical offices and supreme audit institutions. As with evaluation policies, some countries have successfully conducted decentralized evaluations without a central coordination function (e.g., Canada). However, such units are likely to become increasingly necessary given the holistic assessments expected under the SDGs.

Pop-up text for an example of dedicated central evaluation unit: In Colombia, the national evaluation function is driven by a unit located within the Federal Department of Planning, which is responsible for leading the monitoring and evaluation of national development plans. To

do so, it operates SINERGIA (*Sistema Nacional de Evaluación de Gestión y Resultados*), a data collection system that tracks and aggregates information on key national performance indicators. It also manages and reports on 15-20 national priority programmes a year. Although individual departments and subnational administrations also have or are striving to develop evaluation systems and capacities, SINERGIA is central to the government's public-sector performance monitoring effort. The Department has also taken on the role of developing and building decentralized monitoring and evaluation capacities.

3. What human resources will the evaluation entity have?

   a) Human resources will be fully established commensurate with the scale of the national development programme.
   b) The evaluation entity will start with a small team and build gradually.
   c) The evaluation entity will have a small team under a multifunctional division.

4. What professional staff will the evaluation entity have and what will be their role?

   a) The evaluation entity will have a full complement of professional staff to conduct/lead evaluations.
   b) The evaluation entity will have a small number of professional staff to manage evaluations.
   c) Professional staff is not a priority for the evaluation entity.

5. What will be the emphasis of the evaluation function?

   a) Accountability and possibly learning.
   b) Learning.
   c) No specific emphasis.

6. Regarding the level of independence of the evaluation entity, who will it report to?

   a) The evaluation entity will report to the parliament.
   b) The evaluation entity will report to an independent oversight structure.
   c) The evaluation entity will report to a management structure.

Pop-up note on independence: Independent and impartial evaluation processes ensure the evaluation function is independent of other management functions and lends credibility to evaluations.

7. How will the national evaluation policy provide direction on intergovernmental coordination for evaluations?
   a) It will outline interministerial coordination mechanisms, roles and responsibilities.
   b) It will establish coordination committees as and when needed.
   c) It will not address coordination as an immediate priority.

8. How will financial resources for the evaluation function be allocated?

   a) Resources will be annually budgeted.
   b) Resources will be part of the overall programme budget.
   c) Resources will be allocated as needed, with no specific budget assigned.

9. What will financial resources for evaluation cover?

   a) All evaluations in the evaluation plan.
   b) Some of the planned evaluations.
   c) There will be no specific evaluation budget, and resources will be assigned as needed.

10. Who will approve the budget for the evaluation function and evaluations?

    a) Parliament.
    b) An independent oversight structure.
    c) Executive function.

Pop-up note on 'What is a sufficient budget?': Determining a sufficient budget for evaluation depends on the selected parameters, such as the proportion allocated relative to total development spending, assuming emphasis on evaluation of key development areas, and the institutional framework for evaluations. To some extent the budget also depends on the stage of development of the evaluation function.

11. How will evaluations be planned?

    a) An evaluation plan will be prepared annually or biannually.
    b) Evaluations will be planned as part of the key activities of the government.

c) There is no specific planning for evaluations.

12. While making evaluation choices, the following parameters should be taken into consideration. (Select statements that identify the level of priority):

a) Level of impact on policymaking (degree to which evaluation will contribute to a policy action that is critical for improvements in quality of life).
b) Level of urgency for policy inputs (degree to which evaluation will contribute to filling a gap in inputs for policymaking).
a) Level of accountability (entities at different levels of government that would take responsibility for the evaluation results).
c) Net systemic contribution (extent to which evaluation will contribute to a policy action over time that will support the integrated achievements of critical success factors).
d) Level of availability of data (baseline and performance data).
e) Availability of resources (if funds are already assigned for the evaluation).

Pop-up note for evaluation planning: Evaluations must be undertaken in a timely manner so they can inform ongoing decision-making. Planning for evaluation must be an explicit part of planning and budgeting of the evaluation function and the organization as a whole. Annual or multi-year evaluation work programmes should be made public. To ensure maximum utility, plan preparations should include adequate consultations with stakeholders — especially the intended users.

The evaluation plan can be the result of a cyclical or purposive selection of evaluation topics in critical areas of the national development plan. The purpose, nature and scope of evaluation must be clear to all stakeholders. The plan for conducting each evaluation must ensure due process to ascertain timely completion and consideration of the most cost-effective way to obtain and analyse the necessary information. For evaluations to be more useful, the evaluation plans should be integrated with the programme budget cycle.

The plan should be supported with adequate human and financial resources to ensure evaluation quality.

Note: An evaluation plan should include the following:

- Outline of the scope of the plan.
- Purpose and utility of evaluations outlined.

- Whether evaluations are aligned with the planning cycle so they can inform the policy decisions. Whether evaluations are aligned with the planning cycle (so they can inform policy decisions).
- How evaluations will be organized, budgeted and managed (including what resources are needed and where they will come from).
- The practicality of evaluations proposed; process used for prioritizing evaluations.
- Utility of evaluations.
- The practicality of evaluations proposed; process used for prioritizing evaluations.
- Mechanism for executive unit to review and endorse the evaluation plan.
- Scope for responding to ad hoc requests for evaluations not included in the initial plan. (Establish clear guidelines to manage such requests)
- Incorporation of gender equality issues.
- Incorporation of human rights issues.
- Disclosure parameters.
- Dissemination of findings and recommendations.
- Follow-up on evaluations (identify who will decide if recommendations are accepted or not).

13. What provisions will be made to ensure evaluation reports are produced?

   a) All evaluations will include a comprehensive report
   b) Some evaluations will include a comprehensive report
   c) There will be no provision for a comprehensive evaluation report

14. Will funds be assigned to ensure there are adequate resources to produce evaluation reports?

   a) Funds will be assigned to produce comprehensive reports for all evaluations.
   b) Funds will be assigned to produce comprehensive reports for some evaluations.
   c) Funds for a comprehensive evaluation report will not be a priority.

15. What provisions will be made for quality assurance of implementation of the evaluation plan and the evaluation reports produced?

   a) A quality assurance system will be put in place for implementation of the evaluation plan and evaluation reports.
   b) There will be an ad hoc system for quality assurance of the implementation of the evaluation plan and evaluation reports.
   c) There will be no system for quality assurance.

16. What provisions will be made to ensure that evaluation reports are of high quality?

    a) All evaluations will be quality assured by an intergovernmental body or head of the entity.

    b) Selected evaluations will be quality assured by an intergovernmental body or head of the entity.

    c) No provisions will be made for quality assurance / evaluation reports will not be produced.

17. Will specific efforts be made to incorporate gender dimensions in evaluations?

    a) All evaluations will incorporate gender parameters.
    b) Some evaluations will incorporate gender parameters.
    c) There will not be much emphasis on incorporating gender parameters.

18. Will specific efforts be made to incorporate human rights dimensions in evaluations?

    a) All evaluations will incorporate human rights parameters.
    b) Some evaluations will incorporate human rights parameters.
    c) There will not be much emphasis on incorporating human rights parameters.

    d) Pop-up note for evaluation reports: Evaluation reports must present evidence-based findings, conclusions and recommendations in a complete and balanced way. The reports should cover what was found and on what evidence base; what was concluded from the findings relative to the main evaluation questions asked and how conclusions were drawn; what is recommended; and what lessons could be learned from the evaluation, if any. Reports must explain the methodology followed and highlight the methodological limitations of the evaluation and key concerns. They must include a summary that encapsulates the essence of the information contained in the report and facilitates dissemination and distillation of lessons.
    e)
    f) Evidence should include different perspectives and contrary views.
    g)
    h) Evaluation reports must be brief, to the point and easy to understand. They should be presented in a way that allows intended readers to access relevant information in the clearest and simplest manner. Reports should not be overloaded with information that is not directly relevant to the overall analysis.
    i)

19. What provisions will be made to ensure transparency and public sharing of evaluation reports?

    a) Transparent disclosure measures will be in place, and all evaluation reports and findings will be made public.
    b) Evaluation reports and findings will be selectively made public.
    c) Evaluation reports and findings will not be made public.

20. What provisions will be made to ensure the evaluation plan and evaluations are shared?

    a) A full-disclosure policy will ensure that all evaluation-related documents are shared internally and externally (made public) except for those on sensitive subject matters.
    b) There will be only selective external disclosure of evaluation-related documents (excluding those related to sensitive issues), and all evaluations will be disclosed only internally.
    c) There will be no external disclosure of evaluation-related documents.

Pop-up note for disclosure policy: Government entities should have an explicit disclosure policy for evaluations to bolster government's public accountability. Key evaluation products (including evaluation plans, terms of reference, evaluation reports and management responses) should be publicly accessible. A disclosure policy should ensure that the public has easy access to evaluation reports. This requirement is fundamental to fulfilling evaluations' public accountability purpose. Some evaluations may require an exception to the disclosure rule (e.g., when protection of stakeholders' private information is required or unrest could be triggered). In such cases, the disclosure policy should specify the conditions under which an exception can be granted. These conditions should be agreed in advance in a transparent manner.

21. Will mechanisms be developed to ensure use of evaluations?

    a) Mechanisms will be developed to use evaluations for accountability and learning.
    b) Mechanisms will be developed to use evaluations for learning.
    c) Evaluations will be used as needed.

22. How will dissemination of key findings and lessons be handled?

    a) A dissemination strategy for all evaluations is mandatory.
    b) Dissemination is not mandatory and will be based on needs.
    c) Dissemination is not a priority.

Pop-up note: Accountability and learning are closely related, and irrespective of the priorities stated in the evaluation policy there should be sufficient measures in place for evaluations to fulfil this role. The heads of the entities and evaluation functions are responsible for ensuring a mechanism is in place for distilling and disseminating lessons to improve and inform policies and make systemic improvement. Evaluation contributes to knowledge building and organizational improvement. Lessons drawn from evaluations should be accessible to target audiences in a user-friendly way. A repository of evaluations could be used to distil lessons that contribute to peer learning and the development of structured material for informing policy.

C. Mainstreaming evaluation at the ministerial, sector and subnational level

Mainstreaming evaluation in national and subnational institutions consists of fully integrating evaluation activities into all relevant aspects of institutional practice. This includes (a) updating mandates and policies, (b) assigning responsibilities and resources for evaluation, (c) assigning responsibilities and resources for evaluation capacity development, and (d) assigning responsibilities for the use, dissemination and follow-up of evaluation recommendations.

Ideally, where resources are plentiful, officials of various national government departments and provincial and municipal entities become familiar with and regularly conduct or commission evaluations. Where resources are lacking, however, mainstreaming can overburden decentralized institutions with additional tasks without providing the resources to fulfil them. Evaluation responsibilities should therefore be phased in over time, be adapted to the local context and be accompanied by sufficient resources.

Skip to next section (Hyperlink) if mainstreaming evaluation is not a priority.

1. To what degree is evaluation mainstreamed into national institutions?

   a) Evaluation is mainstreamed into all ministries and programmes.

   b) Evaluation is mainstreamed into selected ministries and programmes.

   c) Evaluation is not mainstreamed.

2. To what degree is evaluation mainstreamed into subnational institutions?

   a) Evaluation is mainstreamed into all subnational entities and programmes.

   b) Evaluation is mainstreamed into selected subnational entities and programmes.

   c) Evaluation is not mainstreamed at the subnational level.

3. Will resources be provided for mainstreaming the evaluation function?

   a) Sufficient resources will be assigned to mainstream the evaluation function.

   b) Resources will be available for selected mainstreaming.

   c) No resources will be available for mainstreaming.

4. Will external support be provided for mainstreaming evaluation?

   a) External support is not needed for mainstreaming evaluation.

   b) Partial support in selected areas is needed for mainstreaming evaluation.

   c) Evaluation is not mainstreamed.

   If there is a need for external support:

   - Determine areas where support would be needed.

   - Map possible partners (national institutes, national NGOs, bilateral exchange, donors, international intuitions, international NGOs)

The pop-up text of examples of mainstreaming evaluation: UNESCO has developed a framework for decentralized learning called Learning Cities, which builds on evaluation and other feedback mechanisms to support sustainable municipal development. The framework has been

adapted by the Australian Centre of Excellence for Local Government as a <u>Learning Community Framework and Measuring Impact Toolkit</u> for local governments.

D. Promoting core development values in evaluations

1. Will gender equality issues be integrated into evaluation policy and practice?

   a) Gender equality issues will be fully integrated into evaluation policy and practice.
   b) Gender equality issues will be party integrated into evaluation policy and practice.
   c) Integration of gender equality issues into evaluation policy and practice is not an immediate priority.

Pop-up text: UN Women (2015) Repository of tools and resources for gender mainstreaming. Available at: http://www.unwomen.org/~/media/headquarters/attachments/sections/how%20we%20work/unsystemcoordination/resources%20and%20tools%20for%20capacity%20development%20on%20gender%20mainstreaming%20within%20the%20un%20system.pdf?v=1&d=20150520T181132

2. Will human rights issues be integrated into evaluation policy and practice?

   a) Human rights issues will be fully integrated into evaluation policy and practice.
   b) Human rights issues will be partly integrated into evaluation policy and practice.
   c) Integration of human rights issues into evaluation policy and practice is not an immediate priority.

Pop-up text: OHCHR, WHO, SIDA (2011). 'Human Rights and Gender Equality in Health Sector Strategies: How to Assess Policy Coherence'. Available at: http://www.ohchr.org/Documents/Publications/HRandGenderEqualityinHealthSectorStrategies.pdf

E. Developing a national evaluation community of practice

Skilled government staff and evaluators are needed to commission and conduct evaluations. Providing training to government staff is a useful way to update institutional knowledge. A more sustainable approach, however, involves developing a national evaluation community of practice. A community of practice is a network of individuals and organizations that participate in a field of shared norms and activities. Communities of practice include evaluation practitioners inside and outside the government.

Together, government actors and civil society actors can contribute to the development of a community of evaluators by (a) maintaining an up-to-date framework of evaluation competencies or professional standards; (b) partnering with public and private sector educational institutions to provide short-term and long-term professional development opportunities; and (c) supporting local, national and international knowledge-sharing opportunities, such as journals, conferences, workshops and virtual working groups.

> Pop-up text on examples of civil society associations: At the country level, voluntary organizations for professional evaluators (VOPEs) are being established across countries. (See the list of countries where VOPEs are established.) The Canadian Evaluation Society, one of the oldest VOPEs, maintains and regularly updates a framework of Competencies for Canadian Evaluation Practice. The five competency domains are (a) reflective practice, (b) technical practice, (c) situational practice, (d) management practice and (e) interpersonal practice.
>
> Global evaluation coalitions have also emerged in recent years, such as EvalPartners and the Global Forum of Parliamentarians for Evaluation.

1. To strengthen national evaluations, will partnerships be established with the following?
   a) National public and private sector educational/research institutions: (a) Fully; (b) in some areas; (c) no partnerships envisaged.

   b) National private think tanks: (a) Fully; (b) in some areas; (c) no partnerships envisaged.

   c) National NGOs: (a) Fully; (b) in some areas; (c) no partnerships envisaged.

2. Will there be engagement with a national evaluation community of practice?
   a) They will be engaged generally.
   b) They will be engaged selectively.
   c) There will be no engagement.

3. Will professional development measures be undertaken?

    a) Specific measures will be established to ensure professional development.

    b) Professional development will be undertaken as needed.

    c) Professional development is not an immediate priority.

Summary after first diagnostic process: Briefly summarize the status of evaluation systems, processes and practices; and outline the key enhancements made to the evaluation function during the evaluation capacity diagnostic process.

F. Strengthening data and statistics

Strong data systems are essential to monitoring implementation and impact, conducting evaluation and making evidence-driven policy decisions. Establishing such systems requires having capacities to collect, analyse, store, disseminate and use data. These capacities are usually developed under the mandate of national statistical offices and departments, in line with the United Nations Statistical Commission Fundamental Principles for Statistics, and thus will not be addressed in this document. However, in recent years governments have supported several initiatives to strengthen data and statistical capacities in the light of the SDGs, such as the Asia-Pacific SDG Partnership (2017), the Latin American Data República (2017) and INFRALATAM (2016), the Africa Open Data Conference (2017), and the UN World Data Forum (2017).

Various guidance documents have been produced to support countries in establishing a monitoring framework for the SDGs, including the Sustainable Development Solutions Network (SDSN) report (2015), 'Indicators and a Monitoring Framework for the Sustainable Development Goals: Launching a data revolution for the SDGs', the SDSN SDG Guide (2015), and the World Bank's World Development Indicators Database.

Steps for taking stock of the data needs for evaluations:

- Identify areas where there the data are robust, sufficient or sparse.
- Identify SDG areas that need to be prioritized for data strengthening.
- Take measures and allocate resources to collect data in areas where the statistics are sparse.

52

1. To strengthen connections between evaluation and data systems:

   a) Strong links will be established with statistical institutions.
   b) There will be case-by-case engagement with statistical institutions.
   c) Strengthening connections between evaluation and data systems is not a priority.

2. Will resources be allocated to bridge data gaps?
   a) Resources will be allocated to bridge data gaps for all evaluations.
   b) Resources will be allocated to bridge data gaps case by case.
   c) Bridging data gaps is not a priority.

3. Will partnerships be established to bridge data gaps?

   a) Collaborations will be established to bridge data gaps for all evaluations.
   b) Collaborations will be established to bridge data gaps case by case.
   c) Collaborations to bridge data gaps are not an immediate priority.

4. Will innovations in data and statistics collection for evaluations be explored?

   a) Innovations will be explored fully.
   b) Innovations will be explored partly.
   c) Innovations are not a priority.

G. Report card on strengthening and institutionalizing evaluation capacities

Evaluation capacity diagnostic processes should be conducted periodically. Countries can conduct their own diagnostic exercises, or hire consultants, or participate in regional peer review exercises. These exercises can assess the government as a whole or individual entities. In both cases, they serve to identify needs and orient government action.

Completing the report card below will aid in assessing progress on parameters identified for strengthening and institutionalizing evaluation.

Report card on strengthening and institutionalizing evaluation capacities		
Parameter	Score: 3 is high, 1 is low, 0 is no activity	
	Before diagnostic process	After changes are made
A. Assessing the policy environment.		
B. Strengthening evaluation policy and function.		
C. Mainstreaming evaluation at the ministerial, sector and subnational level.		
D. Promoting core development values in evaluations.		
E. Developing a national evaluation community of practice.		
F. Strengthening data and statistics.		
G. Conducting periodic evaluation capacity diagnostics		
Total score		
Changes in the score		

Changes in scores for the selected period indicate progress as follows. (It is likely some parameters cannot be compared, and overall scores will be adjusted automatically in such cases.)

Advanced: The mechanisms and instruments needed to characterize the evaluation systems are present and entirely adequate to meet decision-makers' needs for evaluative knowledge: >80%
Improved: The mechanisms and instruments needed to characterize the evaluation systems are present, sufficient and satisfactorily meet decision-makers' needs for evaluative knowledge: 70%-<80%
Intermediate: The mechanisms and instruments needed to characterize the evaluation systems are present, partially sufficient and regularly meet decision-makers' needs for evaluative knowledge: 60%-<70%
Incipient: The mechanisms and instruments needed to characterize the evaluation systems are present, but they are partially sufficient and do not regularly meet decision-makers' needs for evaluative knowledge: 50%- <60%
Not structured: The mechanisms and instruments needed to characterize the evaluation systems are partially present, but they are insufficient and do not regularly meet decision-makers' needs for evaluative knowledge: <50%
Percentage improvement in the score

Good: Over 50% increase
Medium: 30%-50% increase
Poor: Below 30% increase

Summary of strengthening and institutionalizing capacities for national evaluation function:

A. Ensure there is adequate supporting data for the responses provided, and summarize key data here.

B. Cross check with relevant actors and provide their perspectives, both supporting and those that are contrary.

C. Provide a brief note on key actors consulted in assessing institutional context.

Evaluation function at the ministerial, sector, programme or subnational level

This section should be completed by entities developing evaluation systems and processes at the decentralized level (ministry, sector, national programme or subnational level). The following sections are adapted for decentralized evaluation function.

A.   Assessing the policy environment.

B.   Strengthening the evaluation framework.

C.   Mainstreaming evaluations at the ministerial, sector, state and local levels.

D.   Promoting core development values in evaluations.

E.   Developing an evaluation community of practice.

F.   Strengthening data and statistics.

A.   Assessing the policy environment

This section aims to determine whether the policy environment is right for carrying out the diagnostic process and developing evaluation policy and practices.

1.   Are there legal issues in developing evaluation policy and practices? (a) There are no legal issues; (b) there are some possible legal issues, but they can be addressed; (c) legal issues remain a constraint and will take some time to be addressed.

2.   Are there constraints on public disclosure of information? (a) There are no constraints; (b) there are constraints on certain subjects; (c) the entity/programme deals with sensitive subject matter requiring a high level of confidentiality.

3.   Is there political resistance to a more structured evaluation practice? (a) There is not any political resistance; (b) there may be some political resistance, but it can be addressed; (c) political resistance remains a constraint and will take some time to be addressed.

4. Is there political space for using evaluation evidence constructively? (a) Significant political space; (b) some political space; (c) no political space.

5. Are there resource constraints in developing evaluation policy and practices? (a) There are no resource constraints; (b) there are some resource constraints, but they can be addressed; (c) resource constraints remain and will take some time to be addressed. .

6. Is there sufficient technical capacity to develop evaluation policy and practices? Technical capacity is: (a) not an issue; (b) partly an issue; (c) a major issue.

B. Strengthening the evaluation framework

Following are a set of parameters for developing an evaluation framework for the entity/programme. The options you choose will be the basis for assessing progress on each of the parameters.

1. How does the evaluation framework guidance apply to the evaluation activities of the entity/programme?

   a) The framework guidance applies to all evaluation activities of the entity/programme.
   b) The framework guidance applies to some activities of the entity/programme.
   c) The framework guidance will be applied to activities case by case.

Pop-up text: Checklist when preparing an evaluation framework for the entity:
- Clarify purpose of the evaluations (accountability and learning or accountability alone).
- Outline the concept and role of evaluation.
- Establish evaluation and entity planning and budgeting links.
- Locate evaluation within an institutional framework.
- Align the entity's evaluation framework with the national evaluation policy and framework (if there is one).
- Establish reporting lines.
- Outline human and financial resources needed.
- Establish accountability and learning processes.

Pop-up text for evaluation framework: Developing an evaluation framework is a key step in building entities' evaluation capacities. The framework should provide a clear explanation of the concept, role and use of evaluation within the entity/programme, including the institutional framework and definition of

roles and responsibilities; links with other entities and central planning processes; an explanation of how the evaluation function and evaluations are planned, managed and budgeted; and a clear statement on disclosure and dissemination.

An evaluation framework is usually a written document that describes rules and responsibilities for evaluation. It is formally endorsed by relevant decision-makers of the entity/programme. An evaluation framework should stipulate what evaluation is and is for; what should be evaluated; how evaluations should be conducted; the need for resources for evaluation; etc. An evaluation framework clarifies expectations around evaluation. It is public and can be used to hold actors to account or to defend evaluations that are conducted, and it informs policy decision processes. The framework also legitimizes building an evaluation culture.

2. Who will lead and coordinate the evaluation function/programme of the entity?

   a) A dedicated, independent evaluation unit.
   b) Personnel in a dedicated evaluation unit within a multifunctional division.
   c) No commitments at this stage.

<u>If the response is 'C' skip to question 5.</u>

Pop-up text on establishing a dedicated central evaluation unit: A central evaluation unit has the mandate to coordinate and oversee evaluation across all areas of government. Its mandate also includes (a) ensuring that decentralized units fulfil their evaluation duties; (b) gathering data and reports from decentralized M&E activities in one place; (c) preparing regular synthesis reports for the executive or legislature; (d) conducting its own national-level evaluations; and (e) contributing to the development of decentralized evaluation capacity. Establishing a central evaluation unit clarifies the national evaluation architecture and creates an institution that parallels national statistical offices and supreme audit institutions. As with evaluation policies, some countries have successfully conducted decentralized evaluations without a central coordination function (e.g. Canada). However, such units are likely to become increasingly necessary in view of the holistic assessments expected under the SDGs.

Pop-up text for an example of dedicated central evaluation unit: In Colombia, the national evaluation function is driven by a unit in the Federal Department of Planning. The unit is responsible for leading the monitoring and evaluation of national development plans. To do so, it operates <u>SINERGIA</u> (*Sistema Nacional de Evaluación de Gestión y Resultados*), a data collection system that tracks and aggregates information on key national performance indicators. The unit also manages and reports on 15-20 priority programmes a year. Although individual departments and subnational administrations also have or are striving to develop evaluation systems and capacities, SINERGIA is central to the government's public-sector performance monitoring effort. The Federal Department of Planning has also taken on the role of developing and building decentralized monitoring and evaluation capacities.

3. What human resources will the evaluation unit have?

   a) The unit will be staffed commensurate with the scale of the development programme.
   b) The unit will start with a small team and build gradually.
   c) The unit will have a small team under a multifunctional division.

4. What professional staff will the evaluation entity have and what will be their role?

   a) The evaluation entity will have a full complement of professional staff to conduct/lead evaluations.
   b) The evaluation entity will have a small number of professional staff to manage evaluations.
   c) Professional staff is not a priority for the evaluation entity.

5. What will be the emphasis of the evaluation function?

   a) Accountability and possibly learning
   b) Learning
   c) There will be no specific emphasis.

6. Regarding the independence of the evaluation function, who will it report to?

   a) The evaluation function will report to parliament/legislative entity.
   b) The evaluation function will report to an independent oversight structure.
   c) The evaluation function will report to a management structure.

Pop-up note on independence: Independent and impartial evaluation processes ensure the evaluation function is independent of other management functions and lend credibility to evaluations.

7. How will the national/subnational evaluation policy address intergovernmental coordination?

   a) The policy will outline interministerial coordination mechanisms, roles and responsibilities.

b) The policy will establish coordination committees as and when needed.

c) Coordination is not an immediate priority.

8. How will financial resources be provided for evaluation?

   a) Resources will be budgeted annually.
   b) Resources will be part of the overall programme budget.
   c) Resources will be located as needed and no specific budget will be assigned.

9. What level of financial resources will be provided for evaluation?

   a) Resources will be provided for all evaluations in the evaluation plan.
   b) Resources will be provided for some of the planned evaluations.
   c) No specific budget will be assigned, and resources will be assigned as needed.

10. Who will approve the budget for the evaluation function and evaluations?
    a) Parliament/legislative entity.
    b) Independent oversight structure.
    c) Executive structure.

Pop-up note on 'What is a sufficient budget?': Determining a sufficient budget for evaluation depends on parameters chosen, for example, the proportion to be spent on evaluation relative to total development spending, what development areas are to be emphasized for evaluation, the institutional framework for evaluations. To a certain extent the budget depends on the evolutionary stage of the evaluation function.

11. How often will evaluation planning take place?

    a) An evaluation plan will be prepared annually or biannually.
    b) An evaluation plan will be prepared as part of the key activities of the government.
    c) There will be no specific planning for evaluations.

12. While making evaluation choices, the following parameters should be taken into consideration. (Select statements that match the priority of the programme/entity):

   b) The level of impact on policymaking (the degree to which evaluation will contribute to a policy action that is critical for improvements in quality of life).
   c) The level of urgency for policy inputs (the degree to which evaluation will contribute to filling a gap in inputs for policymaking).
   d) The level of accountability (entities at different levels of government that would take responsibility for the evaluation results).
   e) The contribution of evaluation (the extent to which evaluation will contribute to changes in development policies and achievement of critical success factors over time).
   f) The availability of data (baseline and performance data availability).
   g) Availability of resources (if funds are already assigned for the evaluation).

Pop-up note for evaluation planning: Evaluations must be undertaken in a timely manner if they are to inform decision-making with relevant information. Planning for individual evaluations must be an explicit part of planning and budgeting for the evaluation function and/or the organization as a whole. Annual or multi-year evaluation work programmes should be made public. To ensure maximum utility, stakeholders (especially the intended users of the evaluation) should be consulted during plan preparations. The evaluation plan should include periodic evaluation of key development issues or one-off evaluations of specific issues in critical areas of the national development plan.

The purpose, nature and scope of evaluation must be clear to all stakeholders. Evaluation plans must ensure timely completion and consideration of the most cost-effective way to obtain and analyse the necessary information. To increase the usefulness of evaluations, evaluation plans should be integrated with the programme budget cycle. Evaluation plans should be supported with adequate human and financial resources to ensure the quality of evaluations.

Note: The following topics should be addressed when preparing an evaluation plan:

- Scope of the plan.
- Purpose and utility of evaluations outlined.
- Whether evaluations are aligned with the government planning cycle so they can inform policy decisions.
- How evaluations will be organized, budgeted and managed.
- Resources needed and where they come from.
- Practical uses of evaluations proposed and the process for prioritizing evaluations.

- Mechanisms for executive review and endorsement of the evaluation plan.
- Guidelines for managing ad hoc requests for evaluations not included in the initial plan.
- Incorporation of gender equality issues.
- Incorporation of human rights issues.
- Parameters for disclosure of evaluations.
- Dissemination of evaluation findings and recommendations.
- Follow-up guidance for determining who will decide whether or not recommendations are accepted.

13. What provisions are made to ensure evaluation reports are produced?

    a) All evaluations will include a comprehensive report.
    b) Some evaluations will include a comprehensive report.
    c) There will not be much emphasis on evaluation reports.

14. What provisions will be made to ensure adequate resources are assigned to produce evaluation reports?

    a) Funds will be assigned to produce comprehensive evaluation reports for all evaluations.
    b) Some evaluations will be assigned funds to produce a comprehensive report.
    c) There will not be much emphasis on comprehensive reports.

15. What provisions will be made for quality assurance of the implementation of the evaluation plan and evaluation reports?

    a) A quality assurance system will be put in place for ensuring the implementation of the evaluation plan and evaluation reports.
    b) There will be an ad hoc system for ensuring the implementation of the evaluation plan and evaluation reports.
    c) There will be no system for quality assurance.

16. What provisions will be made to ensure the quality of evaluation reports produced?

    a) All evaluations will be quality assured by an intergovernmental body or head of the entity.
    b) Selected evaluations will be quality assured by an intergovernmental body or head of the entity.
    c) There will not be much emphasis on quality assurance / evaluation reports will not be produced.

17. What efforts will be made to incorporate gender dimensions in evaluations?

    a) All evaluations will incorporate gender parameters.
    b) Some evaluations will incorporate gender parameters.
    c) There will not be much emphasis on gender parameters.

18. What efforts will be made to incorporate human rights dimensions in evaluations?

    a) All evaluations will incorporate human rights parameters.
    b) Some evaluations will incorporate human rights parameters.
    c) There will not be much emphasis on human rights parameters.

Pop-up note for evaluation reports: Evaluation reports must present evidence-based findings, conclusions and recommendations in a complete and balanced way. The reports should include what was found and on what evidence base; what was concluded from the findings in relation to the main evaluation questions asked, and how such conclusions were drawn; what recommendations emerged from the evaluation; and what lessons could be learned from the evaluation if any. The reports must explain the methodology followed and highlight the methodological limitations and key concerns. They must have a summary that encapsulates the essence of the information contained in the report and facilitates dissemination and distillation of lessons.

Evidence should include different perspectives and contrary views.

Evaluation reports must be brief, to the point and easy to understand. They should be designed to meet the needs of intended users. Evaluation reports should be presented in a way that allows intended readers to access relevant information in the clearest and simplest manner. Reports should not be overloaded with information that is not directly relevant to the analysis.

19. How will transparency and public sharing of evaluation reports be assured?

    a) Transparent disclosure measures will be in place, and all evaluation reports and findings will be made public.
    b) Evaluation reports and findings will be selectively made public.
    c) Evaluation reports and findings will not be made public.

20. How will the evaluation plan and evaluations be shared?

63

a) A full disclosure policy will be prepared, ensuring that all evaluation-related documents will be shared internally and externally (made public) except for those on sensitive subject matters.

b) There will be selective public disclosure of evaluation-related documents (excluding those related to sensitive issues), and all evaluations will be disclosed only internally.

c) There will be no disclosure of evaluation-related documents externally.

Pop-up note for disclosure policy: Government entities should have an explicit disclosure policy for evaluations to bolster the government's public accountability. Key evaluation products (including evaluation plans, terms of reference, evaluation reports and management responses) should be publicly accessible. A disclosure policy should ensure that the public has easy access to evaluation reports. This requirement is fundamental to fulfilling evaluations' public accountability purpose. Depending on the nature of the subject matter evaluated, some cases may require an exception to the disclosure rule (e.g., when protection of stakeholders' private information is required or the evaluation addresses issues that may trigger unrest). In such cases, the disclosure policy should specify the conditions under which an exception can be granted. These conditions should be agreed in advance in a transparent manner.

21. How will use of evaluations be ensured?
    a) Mechanisms will be developed for using evaluations for accountability and learning.
    b) Mechanisms will be constituted for using evaluations for learning only.
    c) Evaluations will be used as needed.

22. How will key findings and lessons from the evaluation practice be disseminated?

    a) A dissemination strategy for all evaluations is mandatory.
    b) Dissemination is not mandatory and will be needs based.
    c) Dissemination is not a priority.

Pop-up note: Accountability and learning are closely related. Regardless of the priority established in the evaluation policy, sufficient measures should be in place for evaluations to fulfil this role. The heads of the entities and evaluation functions are responsible for ensuring that a mechanism is in place for distilling and disseminating lessons to improve and inform policies, policy learning and systemic improvement. Evaluation contributes to knowledge building and organizational improvement. Lessons drawn from evaluations should be accessible to target audiences in a user-friendly way. A repository of evaluation could be used to distil lessons that contribute to peer learning and the development of structured material for informing policy.

C. Mainstreaming of evaluations at the ministerial, sector, state and local levels

Mainstreaming evaluation across different levels of subnational institutions consists of fully integrating evaluation activities into all relevant aspects of institutional practice. This includes (a) updating mandates and policies, (b) assigning responsibilities and resources for evaluation, (c) assigning responsibilities and resources for evaluation capacity development, and (d) assigning responsibilities for the use, dissemination and follow-up of evaluation recommendations.

Ideally, where resources are plentiful, officials of provincial and municipal levels of government are familiar with and regularly conduct or commission evaluations. Where resources are lacking, however, mainstreaming can overburden decentralized institutions with additional tasks without providing them with the resources to fulfil them. Evaluation responsibilities should, therefore, be phased in over time, be adapted to the local context and be accompanied by a commitment of sufficient resources.

Skip to next section (Hyperlink) if mainstreaming evaluation is not a priority.

1. Where will evaluation be mainstreamed in subnational institutions?

    a) In all subnational entities and programmes.

    b) In selected subnational entities and programmes.

    c) There will be no mainstreaming.

2. Will resources be available to mainstream the evaluation function?

    a) Sufficient resources will be assigned to mainstream the evaluation function.

    b) Resources will be available for selected mainstreaming.

    c) There will be no mainstreaming.

3. Is external support provided for mainstreaming evaluation?

    a) External support is not needed for mainstreaming evaluation.

b) Partial support is needed in selected areas for mainstreaming evaluation.

c) There is no mainstreaming of evaluation.

If there is a need for external support:

- Determine areas where support would be needed

- Map possible partners (national institutes, national NGOs, bilateral exchange, donors, international intuitions, international NGOs)

Pop-up text of examples of mainstreaming evaluation: A framework for decentralized learning developed by UNESCO, called Learning Cities, builds on evaluation and other feedback mechanisms to support sustainable municipal development. The framework has been adapted by the Australian Centre of Excellence for Local Government as a Learning Community Framework and Measuring Impact Toolkit for local governments.

D. Promoting core development values in evaluations

1. How will the evaluation framework and practice address gender equality issues?
   a) They will fully integrate gender equality issues.
   b) They will partly integrate gender equality issues.
   c) Gender equality issues in evaluations are not an immediate priority.

Pop-up text: UN Women (2015) Repository of tools and resources for gender mainstreaming. Available at: http://www.unwomen.org/~/media/headquarters/attachments/sections/how%20we%20work/unsystemcoordination/resources%20and%20tools%20for%20capacity%20development%20on%20gender%20mainstreaming%20within%20the%20un%20system.pdf?v=1&d=20150520T181132

2. How will the evaluation framework and practice address human rights issues?

   a) The evaluation framework and practice will fully integrate human rights issues.

b) The evaluation framework and practice will partly integrate human rights issues.

c) Human rights issues in evaluations are not an immediate priority.

Pop-up text: OHCHR, WHO, SIDA (2011). ;Human Rights and Gender Equality in Health Sector Strategies: How to Assess Policy Coherence; Available at: http://www.ohchr.org/Documents/Publications/HRandGenderEqualityinHealthSectorStrategies.pdf

E. Developing an evaluation community of practice

Skilled government staff and evaluators are needed to commission and conduct evaluations. One approach is to provide them with training. A more sustainable approach, however, is to support the development of a national evaluation community of practice — a network of individuals and organizations that participate in a field of shared norms and activities. Communities of practice include evaluation practitioners inside and outside the government.

Together, government actors and civil society actors can contribute to the development of a community of evaluators by (a) maintaining an up-to-date framework of evaluation competencies or professional standards; (b) partnering with public and private sector educational institutions to provide short-term and long-term professional development opportunities; and (c) supporting local, national and international knowledge-sharing opportunities, such as journals, conferences, workshops and virtual working groups.

Pop-up text on examples of civil society associations: At the country level, voluntary organizations for professional evaluators (VOPEs) are being established (see list of countries where VOPEs are established). The Canadian Evaluation Society, one of the oldest VOPEs, maintains and regularly updates a framework of Competencies for Canadian Evaluation Practice. The five competency domains are (a) reflective practice, (b) technical practice, (c) situational practice, (d) management practice and (e) interpersonal practice.

Global evaluation coalitions have also emerged in recent years, such as EvalPartners and the Global Forum of Parliamentarians for Evaluation.

1. Will partnerships be established with public and private sector educational and research institutions to strengthen national evaluations?

- With national and subnational public and private sector educational and research institutions: (a) Fully; (b) in some areas; (c) no partnerships envisaged.

- With national private/subnational think tanks: (a) Fully; (b) in some areas; (c) no partnerships envisaged.

67

- With national/subnational NGOs: (a) Fully; (b) partly; (c) no partnerships envisaged.

2. Will the there be engement with an evalution community of practice?

   a) They will be engaged generally.
   b) They will be engaged selectively.
   c) There will be no engagement.

3. Will professional development measures for evaluation be established?

   a) Comprehensive measures will be established.
   b) Measures will be established as needed.
   c) Professional development for evaluation is not an immediate priority.

Summary after first diagnostics (Briefly summarize the status of evaluation systems, processes, and practices; and outline the key enhancements made to the evaluation function during the evaluation capacity diagnostics)

F. Strengthening data and statistics

Strong data systems are essential to monitoring implementation and impact, conducting evaluations, and making evidence-driven policy decisions. The capacities needed to collect, analyse, store, disseminate and use data are complex. Such capacities are usually developed under the mandate of national statistical offices and departments, in line with the United Nations Statistical Commission Fundamental Principles for Statistics, and thus are not addressed here. In recent years, governments have supported several initiatives to strengthen data and statistical capacities in the light of the SDGs, such as the Asia-Pacific SDG Partnership (2017), the Latin American Data República (2017) and INFRALATAM (2016), the Africa Open Data Conference (2017), and the UN World Data Forum (2017).

Various guidance documents have been produced to support countries in establishing a monitoring framework for the SDGs, including the SDSN report (2015) 'Indicators and a Monitoring Framework for the Sustainable Development Goals: Launching a data revolution for the SDGs', the SDSN SDG Guide (2015), and the World Bank's World Development Indicators Database.

The steps for taking stock of the data needs for evaluation are:

- Identify areas where the data are robust data, sufficient or sparse.
- Identify areas that need to be prioritized for data strengthening.
- Take measures and allocate resources for collecting data in areas where the statistics are sparse.

1. Will connections be made with statistical institutions to strengthen links between evaluation and data systems?

   a) Strong links will be established with statistical institutions.
   b) Engagement with statistical institutions will take place case by case.
   c) Links with data systems are not a priority.

2. Will measures be taken to bridge data gaps?
   a) Resources will be allocated to bridge data gaps.
   b) Resources will be allocated case by case to bridge data gaps.
   c) Measures to bridge data gaps are not an immediate priority.

3. Will partnerships be established to bridge data gaps?

   a) Partnerships will be established to bridge data gaps.
   b) Partnerships will be established case by case to bridge data gaps.
   c) Partnerships to bridge data gaps are not an immediate priority.

4. Will innovations in data and statistics collection for evaluations be explored**?**

   a) Innovations will be explored fully.
   b) Innovations will be explored partly.
   c) Innovations are not a priority.

G. Report card on decentralized evaluation capacities

Capacity diagnostics of different aspects of evaluation should be conducted periodically. Such exercises serve to identify needs and orient government action. Completing the following report card will support progress on parameters identified for strengthening decentralized evaluation. Entities can conduct their own diagnostic exercises, hire consultants to perform the exercise or participate in regional peer review exercises.

Report card on decentralized evaluation capacities		
Parameter	**Score: 3 is high, 1 is low, 0 is no activity**	
	Before diagnostic process	**After changes are made**
A. Assessing the policy environment.		
B. Strengthening the evaluation framework		
C. Mainstreaming of evaluations at the ministerial, sector, state and local levels.		
D. Promoting core development values in evaluations.		
E. Developing an evaluation community of practice		
F. Strengthening data and statistics.		
Total score:		
Overall changes in the score:		
Overall net changes in scores for the selected period will indicate progress as follows. (It is likely some parameters cannot be compared, and overall scores will be adjuste automatically in such cases.) Advanced: The mechanisms and instruments needed to characterize the evaluation systems are present and entirely adequate to meet the decision-makers' needs for evaluative knowledge: >80% Improved: The mechanisms and instruments needed to characterize the evaluation systems are present, sufficient and satisfactorily meet the decision-makers' needs for evaluative knowledge: 70-<80%		

Intermediate: The mechanisms and instruments needed to characterize the evaluation systems are present, partially sufficient and regularly meet the decision-makers' needs for evaluative knowledge: 60-<70%

Incipient: The mechanisms and instruments needed to characterize the evaluation systems are present, but they are partially sufficient and do not regularly meet the decision-makers' needs for evaluative knowledge: 50-<60%

Not structured: The mechanisms and instruments needed to characterize the evaluation systems are partially present but insufficient and do not regularly meet the decision-makers' needs for evaluative knowledge: <50%

<u>Percentage improvement in the score</u>

Good: Over 50% improvement
Medium: 30-50% improvement
Poor: Below 30% improvement

Pop up text: Another example of evaluation diagnosis: The Federal Republic of Somalia and Brazil both conducted surveys of various government departments. Several countries of the European Union diagnosed their evaluation capacities in recent years, using both online surveys and key stakeholder interviews.

Summary of strengthening and institutionalizing capacities for decentralized evaluation function:

A. Ensure there is adequate supporting data for the responses provided, and summarize key data here.

B. Cross-check with relevant actors and provide their perspectives, both supporting and contrary.

C. Provide a brief note on key actors consulted in assessing institutional context.

8. MODULE 4: INTEGRATING THE SDGS AND SUSTAINABLE DEVELOPMENT SPECIFICITIES INTO EVALUATION PROCESSES AND EVALUATIONS

The focus in this module is not on SDG reporting but on ensuring that evaluations integrate SDG principles and approaches.

Pop-up text on resources for mainstreaming SDGs: UNDPs MAPS http://www.undp.org/content/dam/undp/library/MDG/Post2015-SDG/UNDP-SDG-UNDG-Reference-Guide-UNCTs-2015.pdf. SDG score card: Created by the U.K.-based Stakeholder Forum, the SDG Scorecard enables an assessment of each global SDG target in relation to its applicability, implementability, and transformative potential in a national context (Stakeholder Forum 2015). Link

This module responds to some of the key demands that Agenda 2030 places on national evaluation systems and proposes steps for updating evaluation approaches. After mapping national systems against the 2030 Agenda and building or strengthening the basis of a national evaluation system, the next step consists of updating existing evaluation systems considering the SDGs. Agenda 2030 outlines several key principles for evaluation related to achieving the SDGs. Many of these are areas that address strengthening of evaluation systems and process in general, for example, institutional leadership, rigour in the conduct of evaluations, learning systems and partnerships. Agenda 2030 places renewed emphasis on four key issues that national institutions generally emphasize, although with varied emphasis: universality, resilience, partnerships and equity.

1) **Universality**: Ensuring that national programmes/initiatives are well integrated into a broader, holistic approach to sustainable development. It also addresses synergies or conflicts between development initiatives.

2) **Resilience**: Ensuring that the results of national programmes/initiatives are likely to withstand adverse economic, political, social and environmental shocks.

3) **Partnerships**: Ensuring that national programmes/initiatives engage all the required stakeholders, and the programme approaches facilitate collaboration, synergies and collective impact.

4) **Equity**: Ensuring that design and implementation of national programmes/initiatives respect human rights, work to eliminate hunger and poverty, and focus on helping the most vulnerable people and those furthest behind in terms of development.

This module covers:
   A. Endorsing and promoting SDG approaches (Hyperlink)
   B. Establishing multi-stakeholder and multi-institutional approaches (Hyperlink)
   C. Evaluating horizontal policy coherence (Hyperlink)
   D. Evaluating vertical policy coherence (Hyperlink)
   E. Enhancing resilience of development outcomes (Hyperlink)
   F. Forming partnerships for performance data (Hyperlink)
   G. Establishing systems to manage intersectoral/joint evaluations (Hyperlink)
   H. Promoting inclusiveness
   I. Monitoring, reporting and accountability linkages (Hyperlink)
   J. Countries with special development situations (Hyperlink)

A. Endorsing and promoting SDG approaches

1. Are specific measures being promoted to ensure that evaluation approaches respond to SDG-related programme specificities? (a) Fully promoted; (b) being developed; (c) not being developed.

2. Is there an evaluation planning process to prioritize the goals and targets most appropriate for assessing national development outcomes? (a) Fully in place; (b) being developed; (c) not being developed.

3. Is inclusion of civil society and citizens in the evaluation process being promoted? (a) Fully promoted; (b) partly promoted ; (c) not being promoted.

4. Are measures being promoted to emphasize sustainable development approach in evaluations? (a) Fully promoted; (b) partly promoted; (c) not being promoted

5. Are measures to emphasize sustainable development approaches through evaluations being promoted to all relevant stakeholders? (a) Fully promoted; (b) partly promoted; (c) not being promoted.

6. Do the evaluations assess the integration of the SDGs into national development strategies? (a) Fully assess; (b) partly assess; (c) do not assess.

> Pop-up note: Integrating the SDGs requires that government policies and plans at national, subnational and local levels are responsive to SDG requirements and that budget allocations reflect SDG priorities.

B. Establishing multi-stakeholder and multi-institutional approaches

1. Has the entity taken actions to engage multiple institutions in the evaluation of development plans (in contrast to sectoral approaches)? (a) Fully taken; (b) partly taken; (c) no action taken.

2. Are actions taken to obtain a range of citizen perspectives while collecting evaluation data? (a) Fully taken; (b) partly taken; (c) no action taken.

3. During evaluation processes, are actions taken to include the perspectives of marginalized groups? (a) Fully taken; (b) partly taken; (c) no action taken.

4. Are actions taken to inform citizens about evaluation findings and recommendations? (a) Fully taken; (b) partly taken; (c) no action taken.

> Pop-up note on inclusive approaches: Inclusion of the full diversity of stakeholders means paying specific attention to providing representation to all voices, including women and children, with a focus on marginalized groups and individuals. People living in poverty, indigenous communities and other minorities, persons with disabilities, forcibly displaced and stateless persons, children and young people, migrants and people who are lesbian, gay, bisexual or transgender are some of the groups and individuals who are not necessarily included in policymaking and decision-making processes. Including citizens in policy processes and governance for voice requires a political framework and an organizational and procedural structure. Stakeholder engagement in long-term sustainable development works best if it is organized as a continuous process rather than being conducted ad-hoc or through unrelated one-off engagement exercises at different points of the policy cycle.
>
> Similarly for evaluations, stakeholder engagement should be structured with an articulated framework for participation. A structured consultative process enables stakeholders as well as governments to plan and to assemble evidence, reports and other material to make well-researched contributions at the appropriate time in the policy cycle. Standing institutional arrangements support strengthening of the capacities of civil society representatives over time and development of trusting relationships of support and cooperation.
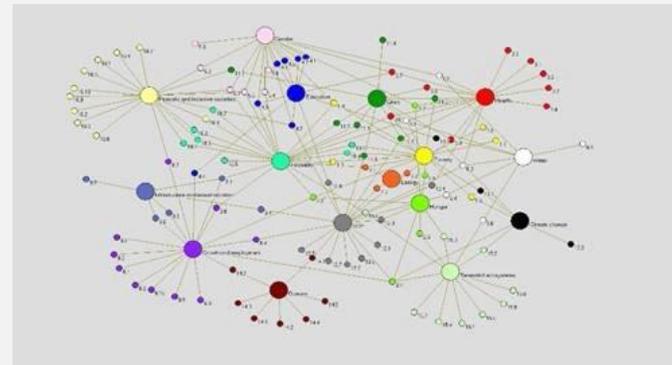
C. Evaluating horizontal policy coherence

Evaluation policies, mandates, competency frameworks, standards and procedures should be updated to clarify new expectations and to ensure the integration of new evaluation approaches, questions and values into evaluation plans, terms of reference and methodologies. Updating these presents an opportunity to foster conversations among evaluation stakeholders about what is needed, desired and feasible.

1. Has the degree of linkage between the SDGs and between the SDG targets been mapped to ensure a more comprehensive evaluation of policy outcomes? (a) Fully mapped; (b) partly mapped; (c) not yet mapped.

2. Have the government institutions involved in achieving a specific programme/target outcome for a more comprehencsive evaluation of policy outcomes been identified? (a) Fully identified; (b) partially identified; (c) not identified.

3. *Has responsibility been assigned to assess how the policies of different departments cohere, contradict and support integrated sustainable development?* (a) Fully assigned; (b) partially assigned; (c) not assigned.

4. *Have sectoral policies been updated to reflect the links between different dimensions of sustainable development?* (a) Fully updated; (b) partially updated; (c) not updated.

5. Have data systems been established to produce disaggregated development data? *(a) Fully established; (b) in progress; (c) not yet established.*

6. Are evaluation approaches updated to take into account issues of policy coherence and integrated programme approaches? (a) *Fully updated;* (b) *partially updated;* (c) *not yet updated.*

Pop-up text for examples and resources on mapping inter-SDG and inter-target linkages:

UNDESA analysis of the SDGs as a network of targets: Using network analysis techniques, UNDESA shows that the SDGs and targets can be seen as a network in which links among goals exist through targets that refer to multiple goals. These links have significant implications for policy integration and coherence across areas. A 2015 UNDESA report notes, "For many of the thematic areas covered by the SDGs, targets relating to those areas are found not only under their namesake goal (when it exists) but across a range of other goals as well. In designing and monitoring their work, agencies concerned with a specific goal (e.g. education, health, economic growth) will have to consider targets that refer to other goals, which, due to the normative clout of the SDGs for development work coming forward, may provide stronger incentives than in the past for cross-sector, integrated work. Similarly, for institutions concerned with monitoring and evaluation of progress under the goals, it will be necessary to look at multiple goals – indeed, all those which include targets

Source: UNDESA-DSD (2015b). 'Towards integration at last? The sustainable development goals as a network of targets'. UN Department of Economic and Social Affairs. DESA Working Paper No. 141. Available at: http://www.un.org/esa/desa/papers/2015/wp141_2015.pdf

referring to one institution's area of interest. This may enable greater integration across goals" (page 9, UNDESA-DSD (2015b). Towards integration at last? The sustainable development goals as a network of targets. UN Department of Economic and Social Affairs. DESA Working Paper No. 141)

- Pajek (Slovene word for 'spider') is a Windows-based program for the analysis of very large networks. UNDESA used this program in its social network analysis of the SDGs and targets. (See Mrvar, A. and V. Batagelj (2015). Pajek, version 3 and 4: Programs for Analysis and Visualization of Very Large Networks – Reference Manual. Available at: http://mrvar.fdv.uni-lj.si/pajek/.)
- Sentinal Visualizer is a program for "advanced link analysis, data visualization, geospatial mapping, and social network analysis (FMS-ASG 2015)." It has been used by the UN Office for Sustainable Development to map the connections among knowledge networks. (FMS-ASG (2015). Sentinel Visualizer: Advanced Link Analysis, Data Visualization, Geospatial Mapping, and Social Network Analysis. FMS Advanced Systems Group. Available at: http://www.fmsasg.com/.)
- 'A Reader's Guide to Social Network Analysis (SNA) Software' provides a website link to a comprehensive listing of network mapping software (Huisman and van Duijn 2011). (Huisman, Mark and van Duijn, Marijtje A.J. (2011). A reader's guide to SNA software. In J. Scott and P.J. Carrington (Eds.) The SAGE Handbook of Social Network Analysis (pp. 578-600). London: SAGE. Available at: http://www.researchgate.net/publication/264896020_A_Reader's_Guide_to_SNA_Software. A listing of software for social network analysis supporting the chapter is available at http://www.gmw.rug.nl/~huisman/sna/software.html)
- Threshold 21 (T21) and iSDG (MI 2015) MI (2011). 'Strengthening Institutional Capacity for Integrated Climate Change Adaptation & Comprehensive National Development Planning in Kenya – Final Report'. Millennium Institute. Available at: https://www.undp-aap.org/sites/undp-aap.org/files/Kenya_Strenghtening%20Institutional%20Capacity%20for%20Integrated%20CCA_july%202011.pd f.
- MI (2015). 'Historical Development and Applications of the T21 Model'. Millennium Institute. Available at: http://www.millennium-institute.org/integrated_planning/tools/T21/timeline.html.

- Howells, et al. (2013). 'Integrated Analysis of Climate Change, Land-use, Energy and Water Strategies'. Nature Climate Change 3, 621–626. Available at: http://www.nature.com/nclimate/journal/v3/n7/full/nclimate1789.html.

- MAMS (World Bank 2015) http://www.worldbank.org/en/research/brief/MAMS ' Integrated Micro-Macro Modeling' (UNDESA 2013) UNDESA (2013). 'Strengthening Macroeconomic and Social Policy Coherence through Integrated Macro-Micro Modelling (2011-2013)'. Available at: http://www.un.org/en/development/desa/policy/capacity/projects_mdgs.shtml.

7. Are formal institutional mechanisms established to enable inter-agency coordination for evaluations?

    a) A coordination committee/commission will be formulated or is already in place.

    b) There will be a designated official with responsibilities for coordination.

    c) Coordination will be performed by existing institutions with cross-sectoral functions (for example planning entity or interministerial bodies) while new entities are developed.

    d) Coordination is not an immediate priority.

8. Is it a priority to identify data sources for evaluating links between the SDGs and between the targets?

    a) Immediate priority.

    b) Intermediate priority.

    c) Not a priority, or data issues will be addressed ad hoc.

Pop-up text for an example of interconnected policies:

Integrated Modelling to Support National Development Planning in Kenya: The Millennium Institute's Threshold 21 (T21) model was applied by the Kenyan Government to "develop more coherent adaptation policies that encourage sustainable development, poverty eradication, and increased well-being of vulnerable groups within the context of Kenya's Vision 2030 program" (MI 2015). The T21-Kenya model was customized to "enable simulations of policies to attain selected MDGs and specific aspects of Kenya Vision 2030 particularly on the economic and social pillars, and used a multi-stakeholder participatory process involving participants from diverse sectors. The T21-Kenya model was used by Kenya's Macro Planning Directorate, Ministry of State for Planning, National Development and Vision 2030, where a core team of 12 modelers was trained to maintain T21-Kenya and use it for policy scenario analysis, with a larger group of 25 government officials also trained in the more general use of System Dynamics and T21".
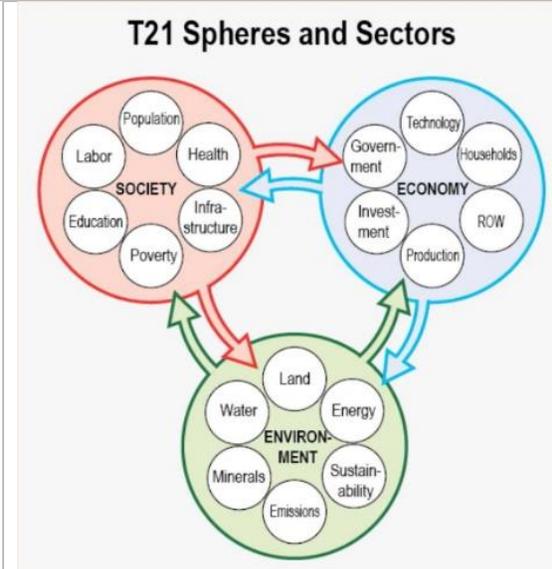
UNDP (2012). Kenya Threshold 21 Dynamic Model Report. United Nations Development Programme – Africa Adaptation Programme. Available at: *https://www.undp-aap.org/resources/publications/kenya-threshold-21-* *dynamic-model-report.*

9. Are institutional processes for assessing integrated policy analysis and intersectoral dimensions of the programmes and policies developed? (a) Fully developed; (b) in progress; (c) the process is stagnant.

10. Are evaluation approaches developed for assessing the responsiveness of programmes and policies to diverse citizen needs? (a) Fully developed; (b) in progress; (c) the process is stagnant.

D. Evaluating vertical policy coherence

Each level of the government is critical in promoting sustainable development and achieving national development outcomes. Establishing links with the private sector and civil society organizations also complements policy coherence. Therefore, integration between planning at different levels of government, between rural and urban areas, and with other complementary actors will be a key focus of national evaluations.

1. Is developing evaluation mechanisms to engage different levels of the government in evaluations a priority? (a) Immediate priority; (b) intermediate priority; (c) not a priority.

2. Is identifying data sources for evaluating policy performance at different levels of the government a priority? (a) Immediate priority; (b) intermediate priority; (c) not a priority / data issues will be addressed as needed.

3. Have measures been taken to bridge performance data gaps at the local level? (a) Fully taken/not an issue; (b) in progress; (c) the process is stagnant.

4. How are formal institutions used to coordinate evaluation across different levels of government? (For example between federal and state governments, state and local governments):

   a) There is a coordination committee/commission.
   b) A designated official has responsibility for coordination.
   c) Existing institutions with cross-sectoral functions are used (for example a planning entity or interministerial bodies) while new ones are developed.
   d) Coordination across different levels of government is not an immediate priority.

   Where relevant:

Are non-state actors (such as civil society and academia) involved in the coordination of evaluation across different levels of government? (a) Fully involved; (b) moderately involved; (c) not involved.

Pop-up text for examples: In Switzerland, the Swiss Federal Office for Sustainable Development leads an array of horizontal and vertical coherence, integration and partnership mechanisms (ESDN 2012). Accountability and implementation of Switzerland's sustainable development strategy use institutional mechanisms for creating both vertical and horizontal coherence, integration and partnerships:

- The **Federal Council** has supreme political responsibility for Switzerland's sustainability policy.
- The Federal Council gives the **Federal Office for Spatial Development** (ARE) the task of coordinating implementation of a sustainability strategy (controlling implementation as well as performing monitoring and evaluation tasks) at the federal level and in collaboration with cantons, municipalities and other stakeholders.
- The **Interdepartmental Sustainable Development Committee** (ISDC) is headed by ARE. This committee furthers the Confederation's sustainable development policy and serves as a platform for sharing information on the Confederation's numerous sustainability activities. Around 30 Swiss government agencies affiliated with ISDC perform tasks relevant to sustainable development.
- In the **Sustainable Development Forum**, ARE works closely with cantons and municipalities and promotes sustainability processes at cantonal, regional and local level.

Source: Taken from MAPS Guidance. ARE (2015). 'Facts and figures. The Swiss Federal Office for Spatial Development (ARE)'. Available at: http://www.are.admin.ch/themen/nachhaltig/00264/00538/index.html?lang=en.
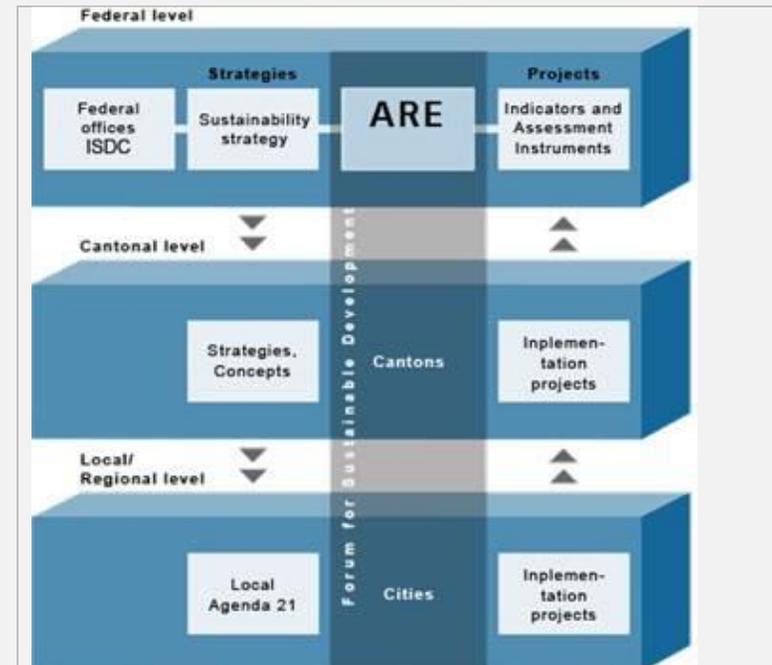
ESDN (2010). 'National Sustainable Development Strategies in Europe: Status Quo and Recent Developments'. ESDN Quarterly Report, September 2010. European Sustainable Development Network. Available at: http://www.sdnetwork.eu/quarterly%20reports/report%20files/pdf/2010-September-National_Sustainable_Development_Strategies_in_Europe.pdf

E. Enhancing resilience of development outcomes

1. Are policy measures for risk management taken in the following categories:

- Economic risk management measures are: (a) fully taken; (b) in progress; (c) not being taken.

- Climate risk management measures are: (a) fully taken; (b) in progress; (c) not being taken.

- Social risk management measures are: (a) fully taken; (b) in progress; (c) not being taken.

2. Is there a shared understanding of what resilience-responsive policy measures entail? (a) There is a shared understanding*; (b) a shared understanding is being developed; (c) there is no shared understanding.
3. Is there a shared understanding of policy measures needed to enhance the resilience of development outcomes? (a) There is a shared understanding; (b) a shared understanding is being developed; (c) there is no shared understanding.

4. Are there indicators in national planning and programming to track resilience? (*a)* There are sufficient indicators; (b) indicators are being developed; (c) there are no indicators.

5. Are there measures to collect data for evaluation of resilience and potential development risks? (a) Such measures are fully in place; (b) such measures are being developed; (c) there are no measures.

F. Forming partnerships for performance data

Forming partnerships can be a cost-effective way to obtain access to new information, skills and technical data. Civil society groups, for instance, may be able to facilitate access to or represent local communities (if the political context allows). They may also have expertise in areas such as gender-based programming and analysis. Evaluators, civil society groups and national statistical offices can not only complement performance data but also ensure that official statistics are robust and disaggregated by key subgroups.

1. Is there collaboration for data on benchmarks and to compare performance? Such collaboration is: (a) fully in place; (b) in progress; (c) stagnant.

2. Are key sustainable development variables that will elicit a fundamental and irreversible change in the behaviour of the system and development outcomes identified for data collaboration? (a) Fully identified; (b) in progress; (c) not identified.

3. Are data for key sustainable development variables that will elicit a fundamental and irreversible change in the behaviour of the system and development outcomes identified for data collaboration? (a) Fully identified; (b) being identified; (c) not identified.

4. Are data sources based on internationally accepted norms and standards identified? (a) Fully identified; (b) being identified; (c) not identified.

Pop-up text on definitions for principles, standards:

- Principles: *A broadly defined and often formally accepted rule;*
- Standards: *Nationally and/or internationally accepted value (i.e., a water quality standard); and*

G. Establishing systems to manage intersectoral/joint evaluations

Ministries and other government entities need to collaborate with one another during the planning and implementation of programmes and policies and in conducting evaluations. An evaluation of progress on child welfare may, for instance, involve ministries of health, gender, education, social affairs, etc. Intersectoral evaluations can be facilitated by a centralized entity that coordinates collection of evaluative information across sectors and conducts its evaluations of progress on national priorities or sets up coordination committees specifically for this purpose.

Pop-up note for examples: South Africa has a centralized evaluation function based in the Department of Planning, Monitoring, and Evaluation (DPME), hosted by the Office of the Presidency. DPME conducts 5 to 10 national-level evaluations a year, which examine priority issues across several departments and levels of administration. In Mexico, CONEVAL (National Council for the Evaluation of Social Development Policy), an autonomous constitutional evaluation entity, evaluates the impact of priority public policies on an integrated, multidimensional poverty index that incorporates indicators on education, health, etc.

1. Are government institutions involved in achieving a specific programme or national development target or outcome identified? (a) Fully identified; (b) partly identified; (c) not identified.

2. Is the level of links between goals and between targets mapped? Such mapping: (a) is completed; (b) is in progress; (c) has not been undertaken.

> Pop-up text for further guidance on how to do the mapping: UNDESA-DSD (2015a). 'Report of the Workshop on Integrated Approaches to Sustainable Development'. UN Department of Economic and Social Affairs. New York, 27-29 May 2015. Available at: https://sustainabledevelopment.un.org/index.php?page=view&type=13&nr=1189&menu=1634.

For countries without a central evaluation entity:

3.   Is it a priority to establish a coordination committee for evaluation of cross-sectoral development issues and to engage diverse stakeholders? (a) High priority; (b) moderate priority; (c) not a priority.

H. Promoting inclusiveness and equity

1.   Do evaluations account for the contribution of actors other than the government? (a) Fully account; (b) partly account; (c) do not account.

2. Do evaluations involve other actors than the government, especially beneficiaries? (a) Fully involve; (b) partly involve; (c) do not involve.

3. Do evaluations assess questions of equity and vulnerability? (*a)* Fully assess; (b) partly assess; (c) do not assess.

4. Are methodologies updated to assess equity issues? *(a)* Fully updated; (b) updating is in progress; (c) not updated.

5. Are there sufficient data and funding to answer questions of equity? (a) Sufficient data; (b) moderately sufficient data; (c) insufficient data.

I. Monitoring, reporting and accountability linkages

"A robust, voluntary, effective, participatory, transparent and integrated follow-up and review framework will make a vital contribution to implementation and will help countries to maximize and track progress in implementing this Agenda to ensure that no one is left behind." – The 2030 Agenda for Sustainable Development.

1. Are links established between development data, monitoring and SDG reporting? (a) Fully established; (b) in progress; (c) not established.

2. *Are links between monitoring and government accountability prioritized?* (a) *Fully prioritized; (b) being prioritized; (c) not prioritized.*

Summary SDG specificities addressed and those that are yet to be considered and reasons for it

J. Countries with special development situations (skip if not applicable)

Some countries in special development situations – such as those facing security fragility or governance fragility; those that are small islands; and least developed countries – might need to evaluate whether the SDG indicator framework is adapted to capture the specificities of their development contexts and needs. Additional indicators, data and monitoring frameworks should respond to their specific needs. For instance, the New Deal for Engagement in the Fragile States and the indicators to monitor Security Council Resolution 1325 might be able to capture the needs and specificities of fragile and conflict-affected countries. Region-specific indicators designed by the UN Economic Commissions and regional intergovernmental institutions might be able to provide solutions for small islands and landlocked and least developed countries.

1. Are evaluations assessing the impact of the development situation on the practice of evaluation? Such assessments are: (a) Fully made; (b) in progress; (c) not being made.

2. Have measures been taken to ensure evaluations pay attention to challenges in special development contexts? (a) Fully taken; (b) in progress; (c) not taken.

3. How are evaluation approaches reformulated to enable assessment under severe data constraints? (a) Evaluation approaches are fully reformulated to address data constraints in special development situations; (b) evaluation approaches are partly reformulated to address data constraints in special development situations; (c) evaluation approaches are not reformulated to address data constraints of special development situations.

Summary of special development situations needs and how evaluations responded to them

K. Report card: Integrating the SDGs and sustainable development specificities in evaluation processes and evaluations

Report card on integrating the SDGs and sustainable development specificities into evaluation processes and evaluations		
Parameter	**Score: 3 is high, 1 is low, 0 is no activity**	
	Before diagnostic process	**After changes are made**
A. Endorsing and promoting the SDG approaches		
B. Establishing multi-stakeholder and multi-institutional approaches		
C. Evaluating horizontal policy coherence		
D. Evaluating vertical policy coherence		
E. Enhancing resilience of development outcomes		
F. Forming partnerships for performance data		
G. Establishing systems to manage intersectoral/joint evaluations		
H. Supporting inclusiveness		
I. Monitoring, reporting and accountability linkages		
J. Countries with special development situation		
Total score		
Overall changes in the score		

Overall net changes in scores for the selected period will indicate progress as follows. (It is likely some parameters cannot be compared, and scores will be adjusted automatically in such cases.)

Score
Advanced: The mechanisms and instruments needed to characterize the evaluation systems are present and entirely adequate to meet decision-makers' needs for evaluative knowledge: >80%
Improved: The mechanisms and instruments needed to characterize the evaluation systems are present, sufficient and satisfactorily meet decision-makers' needs for evaluative knowledge: 70-<80%
Intermediate: The mechanisms and instruments needed to characterize the evaluation systems are present, partially sufficient and regularly meet decision-makers' needs for evaluative knowledge: 60-<70%
Incipient: The mechanisms and instruments needed to characterize the evaluation systems are present and partially sufficient but do not regularly meet decision-makers' needs for evaluative knowledge: 50-<60%
Not structured: The mechanisms and instruments needed to characterize the evaluation systems are partially present but insufficient and they do not regularly meet the decision-makers' needs for evaluative knowledge: <50%

Percentage improvement in the score

Good: Over 50% improvement
Medium: 30-50% improvement
Poor: Below 30% improvement

Summary of measures for integrating SDG specificities in evaluations:

A. Ensure there is adequate supporting data for the responses provided, and summarize key data here.

B. Cross-check with relevant actors and provide their perspectives, both supporting and contrary.

C. Provide a brief note on key actors consulted in assessing institutional context.