THE UNITED REPUBLIC OF TANZANIA

PRIME MINISTER'S OFFICE

POLICY, PARLIAMENT AND COORDINATION

NATIONAL
EVALUATION MANUAL

JUNE, 2024

PREFACE

Tanzania is at the end of the 2025 development vision and has begun the process of developing a new development vision for 2050. This is an important step in the development of the country. However, the new vision to be developed will require a robust monitoring and evaluation system that emphasis on Evaluations and enable monitoring of various development interventions and also to conduct evaluations of the implementation of those developmental interventions.

For a quite long-time, Tanzania has managed the monitoring and Evaluation system by focusing more on monitoring while the evaluation part being conducted very sporadically and thus not helping to properly manage the implementation of Policies, Strategies, Plans, programmes and projects in the country. With this in mind, and given the importance of the Government conducting evaluations of its development interventions, it is time to have specific tool that will provide guidance to Public Service institutions in conducting evaluation of its various development Interventions.

Thus, this Manual is the first of its kind being developed in Tanzania for Public Sector Institutions. Its development, mark the beginning of Tanzania Government to utterly value the importance of M&E in general and create an evaluation culture within Government machinery.

I urge all responsible Public Service accounting officers to facilitate Monitoring and Evaluation Divisions, Units and Sections to ensure fully utilized of this manual to assist evaluations activities of different interventions in our Policies, Strategies, Plans, Programs and Projects. To this end, government and in particular, Division of Performance, Monitoring and Evaluation (DPME) will not only be relying on outsourcing evaluations, but rather focus on capacitating the public Service Institutions to undertake evaluations through training and partnerships. Strategic support is critical to fostering peer learning amongst government officials and the achievement of performance outcomes. It is our goal to promote the institutionalisation of evaluations across the public sector to promote efficient service delivery, and ensure that government interventions have a meaningful impact on beneficiaries.

Hon. Jenista Joachim Mhagama (MP)
Minister of State, Prime Minister's Office
Policy, Parliament and Coordination

FOREWORD

The proposed National Evaluation Manual (NEM) focuses on evaluations of strategic and important Policies, Plans, Programmes or Projects, which are identified as part of a National Evaluation Plan. It further provides a basis for Public Institutions to make a full use of mixed evaluation approaches.

The NEM provides background information including the overview on the existing situation of evaluation practice in Tanzania as well as the purpose, principles, methodologies and techniques for conducting evaluation based on different types of evaluation. Moreover, it outlines the steps to be taken in planning for evaluations and presents ways through which the evaluation process would ensue and be managed including management response and implementation plans. It also present evaluation quality assessment so as to ensure adherence to the predetermined standards and criteria stipulated in evaluation documents.

Its scope extends to all MDAs, Parastatals, Regional Secretariats and LGAs. It also underpinned by a utilization-focused approach to ensure that evaluations are used to improve programme performance, promote accountability, support effective evidence-base decision-making, and promote knowledge creation and dissemination based on the three categories of evaluations, which are Ex-ante, Midterm or Post-Ante Evaluations. Therefore, the type of evaluation chosen is determined by its purpose. The NEM also includes the implementation of priority portfolio of evaluation approaches, namely rapid evaluations, sectoral reviews, Climate, and a gender focussed evaluation etc.

Furthermore, the strategic identification of evaluations necessitates that the selection be aligned to priorities of government as well as key strategic frameworks and plans such as the Vision 2050, Rulling Manifesto, Five Year Development Plans (FYDPs), Medium-Term Strategic Framework (MTEF) and the existing International Planning Frameworks, thereby ensuring synergy between government planning and processes of evaluation of various development interventions.

Evaluation is a fundamental component of the standard operating procedures of all Government Institutions and it form part of the policy or programme cycle. This Manual will assist in cultivating a culture for evaluation and improve performance management. It will enable the government to generate evidence that will contribute to the determination of performance on the Tanzania Development Vision, Five-Year National Development Plans etc. Furthermore, it will ensure adherence to principles of methodological soundness, data, and information management. This will enable evidence-based policy and decision-making, as envisaged in the Monitoring and Evaluation Guideline of 2024 and other planning frameworks.

The Public Service Employment Policy of 1998, recognizes that Monitoring and Evaluation is an integral part of the managerial responsibilities of the MDAs, RSs and LGAs. Monitoring and evaluation questions the rationale of implementing a plan, program or project, and provides information on its performance, results and cost-effectiveness. The findings and recommendations of such evaluations can be used by MDAs, RSs and LGAs to make more informed decisions on the management and resourcing of their programs; to be accountable for the programs for which they are responsible; and to provide quality advice to the Government.

As part of its responsibilities for effective managing Government Performance, the Prime Minister's Office has issued this Evaluation Manual, which covers the establishment and continuing operation of Evaluation in MDAs, RSs and LGAs. The major purpose of this document is to provide useful framework for public institutions conducting sound evaluation. Hence the document identifies and discusses the factors that are useful to consider in an evaluation and helps to identify the important factors to be considered in assessing the quality of evaluations. These factors represent an initial agenda of the items that will be considered by the Prime Minister's Office whenever it comments on the quality of selected evaluations. It is more beneficial to develop explicit standards after the government has had more experience with the conduct and use of evaluation studies.

Typically, Government goes about its work through cycles of planning, budgeting, implementation, monitoring and reporting, and evaluations. Besides its importance in the lifecycle of Government institutions, evaluation can also be used to inform ongoing strategic management and decision-making of policy implementation, programmes and projects. This is a manual and it is not meant to be prescriptive. There is a need for a better monitoring of improvement plans, to ensure that evaluation findings are used to improve service delivery.

Dr. Jim James Yonazi,
Permanent Secretary
Prime Minister's Office – Policy, Parliament and Coordination.

EXECUTIVE SUMMARY

For a quite sometimes, Evaluations in Tanzania have been conducted on an ad-hoc basis across the MDAs, RSs and LGAs without having one common framework standardizing the way evaluations should be undertaken. This National Evaluation Manual (NEM) provides detailed understanding of standards of evaluations practices for the National Evaluation System (NES) in Tanzania Public Service Institutions.

The structure of this Manual has divided into five chapters of which chapter one provides a background information including the overview on the existing situation of evaluation practice in Tanzania as well as the purpose. Chapter two dwells onto the principles, methodologies and techniques for conducting evaluation based on different types of evaluation. Moreover, chapter three outlines the steps to be taken in planning for evaluations while chapter four presents ways through which the evaluation process would ensue and be managed including management response and implementation plans. The last chapter present evaluation quality assessment so as to ensure adherence to the predetermined standards and criteria stipulated in evaluation documents.

Hence, the development of this document aims to provide standardized and systematic procedures for undertaking evaluations at different levels within government. The formulation and use of this Manual will therefore: - Encourage and improve the use of data for informed and evidence decision –making; Enhance and instill M&E culture across the public sector; Strengthen quality of data generated from various source within the public sector; Establish consistence in conducting evaluations; Guide the institutionalization of evaluations systems in the country; Maintain methodological standards and evaluation criteria; Ensure value for money in implementation of Government initiatives; Facilitate quality assurance of the evaluations conducted; Promote a learning culture among institutions; Enhance transparency and accountability during the implementations of the projects, programs and policies, and Ensure evaluation timelines and proper use of evaluation findings.

Its scope covers all Government Interventions carried out in the Ministries, Independent Departments and Regulatory Authorities, Executive Agencies, Parastatals, Regional Secretariats and Local Government Authorities. Moreover, the National Evaluation Manual (NEM), provides principles, methodologies and techniques for managing and conducting Evaluations in Tanzania. It also underpinned by a utilization-focused approach to ensure that evaluations are used to improve programme performance, promote accountability, support effective evidence-base decision-making, and promote knowledge creation and dissemination. These evaluations are conducted at National level (through Countries' Flagship Projects), Sector, and at Institutional levels.

The guideline adopts the African Evaluation Principles as well as the Organization for Economic Cooperation and Development – Development Assistance Committee (OECD-DAC) principles and criteria. In undertaking evaluations in Tanzania, Internal1, External2 or Hybrid3 evaluators should be used. For external evaluators the government procurement system which comply to the Public Procurement Act [CAP. 411 R.E. 2019], Public Finance Act [CAP.348 RE 2020] as well as The Local Government Finance Act. [CAP. 290 R.E.2019], The Public Corporation Act CAP 257 and The Company Act CAP 212 must be adhered to.

Additionally, all Government Institutions are encouraged to triangulate. It helps to generates findings that are hopefully accurate, reliable, credible, unbiased and validated. It also presents ways through which the evaluation process would ensue and be managed by the commissioner of that evaluation. It also presents the management response and implementation plans which is critical to ensuring that the MDAs, RSs and LGAs can utilize the evaluative findings. For the generation of knowledge, results-based management, and accountability to program partners, evaluations of excellent quality are crucial. Finally, evaluation quality assessment is also critical as it offers a critique of an evaluation's approach, the veracity of the findings and evaluative information, and the reliability of the recommendations and conclusions drawn from the methods employed during the evaluation process.

1 *Internal evaluators are the public Servants in the M&E Divisions or Unit in the MDAs, RSs and LGAs*

2 These are independent evaluators who are not employed by the MDAs, RSs and LGAs or any other ministry within the government of the United Republic of Tanzania or any public sector organization whose intervention is being assessed.

3 A hybrid evaluation involves both internal and external staff working together. Hybrid evaluations is seeking to bring together the advantages of both types of evaluators.

Table of Contents

LIST OF ACRONYMS AND ABBREVIATIONS

ASDP II	Agricultural Sector Development Programme II
Atlas ti	Qualitative Data Analysis Software
CBOs	Community-Based Organisations
DED	District Executive Director
DPLO	District Planning Officer
ETRG	Evaluation Technical Reference Group
FYDP	Five-Year Development Plan
GoT	Government of Tanzania
LGAs	Local Government Agencies
LGA-EPs	Local Governments Authorities- Evaluation Plans
MDAs	Ministries, Departments and Agencies
MDA-PP	Monitoring and Evaluation for Development Effectiveness and Accountability Project
MDA-Eps	Ministries Departments and Agencies- Evaluation Plans
M&E	Monitoring and Evaluation
MESA	Monitoring and Evaluation Situation Analysis
NFYDP	National Five-Year Development Plan
NEEC	National Economic Empowerment Council
NEP	National Evaluation Plan
NGO	Non- Government Organisation
NM&E	National Monitoring and Evaluation
NVIVO	Qualitative Data Analysis Software
PO-RALG	Public Office-Regional Administration and Local Governments
PMO	Public Management Office
PMO-PPC	Program Management Office-Planning, Programming, and Coordination
SPSS	Statistical Package for Social Sciences
STATA	Statistical Analysis and Data Management Software
OECD-DAC	Organisation for Economic Cooperation and Development – Development Assistance Committee
PMED	Performance Monitoring and Evaluation Department
POPSMGG	President’s Office Public Service and Good Governance
RFP	Request for Proposal
SIDA	Swedish International Development Cooperation Agency
SMART	Specific, Measurable, Achievable, Relevant and Time-Bound
TOC	Theory of Change
ToR	Terms of Reference
UN	United Nations
USAID	United States Agency for International Development
NSDP	National Skills Development Programme

CHAPTER

1 INTRODUCTION AND BACKGROUND

1.1. Introduction

The goal of the Government of Tanzania (GoT) as stipulated in the Tanzania Vision 2025 is to transform its economy to Semi-industrialized economy and realization of quality service delivery to its citizens. Policies, strategies, and reforms have thus been instituted to ensure the realization of this goal. The efforts to realize the goals call for the Government intervention to establish and strengthen Monitoring and Evaluation System for the purpose of ensuring that every government entity postulate their mandates towards achieving this goal. Thus, in 2014, the first Monitoring and Evaluation Systems Framework for Tanzania Public Service was adopted and propagated. The purpose of the framework was to outline key components of the Tanzanian Monitoring and Evaluation System, Institutional arrangements and roles and responsibility on managing and sustaining M&E system as well as to address the identified challenges in implementing monitoring and evaluation in the country.

These challenges identified includes, absence of a mutual understanding across government of what constitutes monitoring and evaluation as well as integrated monitoring and evaluation systems and processes. The 2014 framework enlists evaluations as a key component of this system. According to the framework, evaluations *"intend to obtain evidence as to whether the interventions implemented and outputs produced have led to achievement of the outcomes as envisioned in the strategic plan, project or programme outputs at Ministries, Departments and Agencies (MDAs), Reginal Secretariat and Local Government Agencies (LGAs), Sector or National levels."*[1] Therefore, evaluations are key for the Government for the following reasons:

- **Strengthening learning and capacity**: Learning is a vital aspect in evaluation. The application of the learning concept in Monitoring and Evaluation (M&E) system within government machinery is important for informed decision-making and program improvement. Its integration into Government framework and guideline enhance effectiveness and efficiency of various intervention, ultimately leading to more impactful outcomes and better resource allocation.

- **Accountability and transparency:** As much as the learning aspect is crucial part for evaluation, MDAs, RSs and LGAs also need to be accountable and transparent to the citizens with respect to projects and their results. Moreover, effective coordination of M&E activities is essential to ensure that the information generated through M&E activities is used to promote accountability and transparency.

1. M&E Systems Framework (2014: 23)

- **Data Demand and Use.** In recent years, the demand for monitoring and evaluation activities in Tanzania has been increasing. This demand emerged after the inception of Tanzania Development Vision 2025. Results from Implemented Plans, Programs and Projects are also demanded in Parliament. This is significant as it prompts the Government to put more effort into evaluations to inform the public. At the same time, the Government needs the evaluation results for decision-making. In this regard, demand of and supply thereof is important for both parties.

1.2 Tanzania Monitoring and Evaluation Situation Analysis

The current environment in Tanzania is conducive for monitoring and evaluation (M&E) as both government and non-government entities are aware of the importance of monitoring and evaluation. This can be evidenced by multiple initiatives taken by the Government to improve M&E System. These initiatives include, the establishment of the Monitoring and Evaluation Systems Framework for Tanzania Public Service of September (2014), the presence of a central institution, Performance Monitoring and Evaluation Division (PMED), in the Prime Minister's Office (PMO) and reestablishment of the Planning Commission in June, 2023. Further, the establishment of fully-fledged M&E unit within all Ministries; inauguration of the national Monitoring, Evaluation and Learning forums; development of national M&E frameworks and tools such as, Project M&E guideline of 2021; M&E guideline for RS and LGA, introduction of Government Coordination Dashboard, and the ongoing efforts to engage national and international stakeholders to strengthen M&E system, signals the value bestowed on M&E in the country.

Despites the aforementioned initiatives and achievements, the Tanzania public sector continues to face challenges in regards to M&E, such as, inadequate monitoring and evaluation skills, inadequate funding, data quality, limited utilization of M&E findings and absence of a national policy on monitoring and evaluation which is a key instrument to structure and institutionalize the practice at both national and local levels within the government structures. Additionally, the implementation of M&E functions/activities within the Government institutions are more on monitoring such as field visits and financial tracking in regards to projects and programs. Hence, evaluations of different typologies are carried out at a meager pace, infrequently, with low quality and the results produced thereof are rarely utilized for improvement of government performance. This is attributed by, lack of managerial and technical guidance on evaluations, Inadequate evaluation technical capacity, Lack of data management skills, evaluations being considered costly, leaving aside the opportunity cost aspect and little understanding and insight on the relevance of evaluations.

Consequently, evaluations for government interventions are not a priority during the institutional planning and budgeting process. Thus, evaluation undertakings are donor driven and focus on programs and projects especially service delivery-based sectors. For example: sectors such as the education, health and water conduct evaluations with an intention of creating evidence for establishing new project or program; a continuation of a new funding of another cycle or phase of the project/program. This implies that, very few evaluations are conducted to the productive sectors including manufacturing, agriculture, tourism and creative industry.

Although, public service officials seem to understand the value of evaluations, they do not have the necessary skills required to undertake effective evaluations. While many officials who undertake M&E activities have a background in social sciences disciplines such as statistics, economics and planning, which is a good base to build evaluation capacity. Other crucial skills are also required to build and sustain a national evaluation system, within other social science discipline.

Additionally, Public service officials are overburdened with reporting and conflicting formats. They report on different templates to different audiences within the public service system which consumes a lot of their time. Also, it erodes their ability to focus on what their organizations are learning from monitoring data. Thus, insights from monitoring activities are not always informing decisions being taken on budget allocations and programme planning. Institutionalization of M&E through policy and government investment in M&E infrastructure can improve these linkages.

1.3 Purpose of the National Evaluation Manual

Evaluations in Tanzania have been conducted on an ad hoc basis across the MDAs, RSs and LGAs without having one common framework standardizing the way evaluations should be undertaken. There have been varied methodological approaches, designs and standards adopted across government institutions. Hence, the development of this document aims to provide standardized and systematic procedures for undertaking evaluations at different levels within government. The formulation and use of this Manual will: -

(i) Encourage and improve the use of data for informed and evidence decision – making;

(ii) Enhance and instill M&E culture across the public sector;

(iii) Strengthen quality of data generated from various source within the public sector;

(iv) Establish consistence in conducting evaluations;

(v) Guide the institutionalization of evaluations systems in the country;

(vi) Maintain methodological standards and evaluation criteria;

(vii) Ensure value for money in implementation of Government initiatives;

(viii) Facilitate quality assurance of the evaluations conducted;

(ix) Promote a learning culture among institutions;

(x) Enhance transparency and accountability during the implementations of the projects, programs and policies, and

(xi) Ensure evaluation timelines and proper use of evaluation findings.

The manual aims to assist the cultivation of culture for evaluation. It will enable the government to generate evidence that will contribute to the determination of performance on the National Development Vision and Five-Year National Development Plans. Furthermore, it will ensure adherence to principles of methodological soundness, data, and information management. This will enable evidence-based policy and decision-making, as envisaged in the Monitoring and Evaluation Systems Framework.

1.4 Structure of the Manual

The document is structured into five chapters of which chapter one provides a background information including the overview on the existing situation of Evaluation practice in Tanzania as well as the purpose. Chapter two dwells on the principles, methodologies and techniques for conducting evaluation based on different types of evaluation. Moreover, chapter three outlines the steps to be taken in planning for evaluations while chapter four presents ways through which the evaluation process would ensue and be managed including management response and implementation plans. The last chapter present evaluation quality assessment so as to ensure adherence to the predetermined standards and criteria stipulated in evaluation documents.

1.5 Scope

The principles outlined in this document concern the evaluations of all intervention by all Government Interventions including carried out in the Ministries, Independent Departments and Ombudsman, Executive Agencies, Parastatals, Government Companies, Regional Secretariats and Local Government Authorities.

CHAPTER

2 PRINCIPLES OF EVALUATION, METHODOLOGIES AND TECHNIQUES

2.1 Introduction

Tanzania Monitoring and Evaluation framework of 2014, recognises Evaluation as the fourth component of its Government of Tanzania (GoT) Monitoring and Evaluation system. This consists of feasibility studies, process, project or programme evaluations, thematic evaluations, performance audits, outcome evaluations, impact evaluations and population census. The evaluations to be conducted by MDAs, RSs and LGAs are usually shown in the Evaluation Plans as part of the Results Framework in their strategic plans, projects, or programme documents. It is expected that the evaluations will form the basis for the future work of the institution and the major findings will be used for decision making and documented in the Performance Reports.

Moreover, the National Monitoring and Evaluation System in the Public Sector helps to enforce the implementation of a government intervention. It also underpinned by a utilisation-focused approach to ensure that evaluations are used to improve programme performance, promote accountability, support effective evidence-base decision-making, and promote knowledge creation and dissemination. These evaluations are conducted at National level (through Countries' Flagship Projects), Sector, and at Institutional levels. This chapter presents the fundamentals of evaluation namely: definition of evaluation, evaluation criteria, types of evaluations and the methods of data collection during evaluations.

2.1.1 Definition of evaluations

Evaluation is defined as the systematic and objective assessment of the design, implementation, and results of ongoing or completed projects, programs and policies and other interventions. Monitoring on the other hand provides real time data which can be used as a basis during evaluation. The table below provides a distinction between the two concepts:

Table 1: Difference between monitoring and evaluation

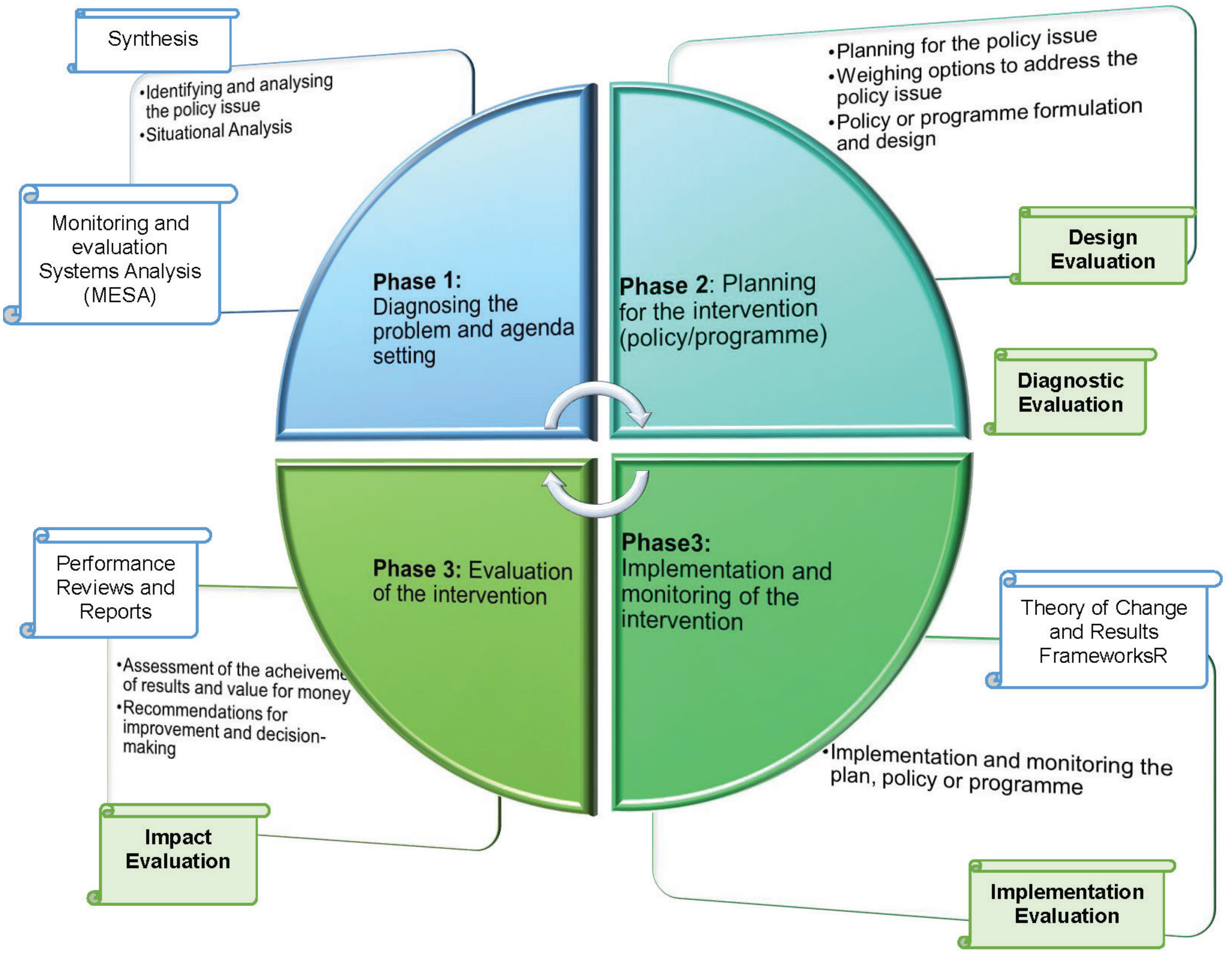
	Monitoring	Evaluation	Audit
Data collection	Routinely collect data on indicators, targets, and actual results	Interprets the collected data and draws conclusions about the findings regarding linkages between targets and actual results	Review documents such as invoices, contracts, and bank statements to formulate an audit opinion based on the evidence gathered and documented.
Data use	Systematizes, classifies, validates, and stores data	Processes, mines, and refines the stored data to extract the most accurate information needed to identify relevant issues or to fill gaps in the data	Accountability
Output	Reports comparisons, differences, and similarities between comparable earlier and later data sets in formats that indicate change over time	Interprets the data, assesses, and makes value judgements about the extent of progress of lack thereof, and to what extent the results are good or bad	Financial Compliance reports
		Analyses how the results were achieved, how this compares to the intended objectives and process, the causal linkages, and makes recommendations for the next steps (termination, adaptation etc.) and how the intervention can be strengthened	Draft a letter to management highlighting any internal control deficiencies and other significant issues found during the audit.

Source: Rabie and Goldman, 2014

Evaluations are an integral part of the policy and programme cycle, facilitating evidence-based decision and policy making. Integrating planning, monitoring, evaluation, and reporting systems promotes accountability and transparency, which are core principles of the public sector. Further, it ensures an effective and efficient implementation of performance management systems.

During the strategic planning process, MDAs, RSs and LGAs should develop a results framework and provide reports on the implementation of the contents of this framework

The following diagram shows the different types of evaluation that can be conducted at each phase of the policy or programme cycle:



2.1.2 Evaluation Criteria

As a result of a common practice that is outlined in the situation analysis, this manual will provide a basis for which public service institutions will conduct evaluations by the use of evaluation criteria as a benchmark to assess interventions.

It adopts the Organisation for Economic Cooperation and Development – Development Assistance Committee (OECD-DAC) and African Evaluation Principles. Evaluation criteria are objective principles which organize the analysis and findings to provide useful information on the extent to which the intervention was relevant, appropriate, efficient, and effective. The application of the DAC criteria is based on two proposed principles discussed below[2]:

Principle 1

To enable high-quality, meaningful evaluation, the criteria should be applied with care. They must be contextualised, in other words, they must be understood in the context of the specific evaluation, the intervention being assessed, and the stakeholders participating. The evaluation questions (what you want to know) and what you want to do with the answers should guide how the criteria is understood and analysed.

Principle 2

The application of the criteria is determined by the goal of the examination. The criteria should not be applied in a mechanical manner instead, it should be addressed in accordance with the needs of the appropriate stakeholders and the context of the evaluation. With respect to the goal of the evaluation, either more or less time and resources may be allocated to the evaluative analysis for each criterion. Data accessibility, resource limits, timeliness, and methodologies may all have an impact on how (and whether) a specific criterion is covered.

The evaluation criteria must be clarified in the evaluation framework developed by the evaluator (appointed or internal). In addition, the framework should also outline the approach to the evaluation, instruments to be used and the sources of information. The table below outlines the main evaluation criteria the objective and key questions for each criterion.

2 https://www.oecd.org/

Table 2: OECD DAC evaluation criteria

Evaluation Criteria	Objective	Questions it responds to

Relevance	Measures the extent to which the programme design meets the need or problem it is intending to address.
	
	Are we doing the right thing?
		- To what extent is the programme aligned to the government priorities?
		- Identify the main problem or need that the programme seeks to address?
		- To what extent is the programme approach or design able to address the policy or programme issue identified?
		- To what extent does the Theory of Change (ToC) outline the causal mechanism for achieving the envisioned outcomes and impacts?
		  o What are the assumptions underpinning the ToC?
		  o Have the indicators for success been defined and are they SMART[3]?
		  o Are there appropriate financial and other measurement systems in place?

Coherence	Examines the programmes compatibility with other interventions in the country, sector, or institution How well does the intervention fit?	- To what extent is the intervention aligned with wider policy frameworks and other interventions of the institution? - To what extent is the intervention aligned with external policy commitments of the country, sector, and institution?

3 Specific, Measurable, Achievable, Relevant and Time-bound

Evaluation Criteria	Objective	Questions it responds to

Efficiency	Measures the extent to which the intervention achieves, or is likely to achieve, its intended results in a timely and economic manner
	
	How well are resources being used?
		• Have the results of the intervention been achieved in a cost-efficient manner?
		• Are there alternatives for achievement of the results with less inputs or funds?
		• Where the results achieved within the intended timeframe?
		  ○ How do the assumptions made in the design of the programme influence the implementation?
		  ○ What are the main enablers and constraints/ barriers to efficient implementation?

Effectiveness	Measures the extent to which outcomes were achieved, or expected Is the intervention achieving its objectives?	• What factors contributed to or prevented the achievement of objectives? • To what extent did the different beneficiary groups find the intervention valuable? • Are beneficiaries able to recognise and identify the changes brought about by the interventions?
Impact	The effects due to implementation of the intervention. This could be intended or unintended effects. It can be primary or secondary effects What difference does the intervention make?	• What has been the social, environmental, and economic effects of implementing the intervention? • How has the intervention affected its beneficiaries? • What are some of the emerging lessons and learning from the programme?

Evaluation Criteria	Objective	Questions it responds to
Sustainability	Examines the extent to which the benefits of the intervention continue or likely to continue Will the benefits last after the programme has been completed?	• Are there any indications that the impacts from the programme are sustainable over the long-term? o If yes, what factors influence the sustainability of the programme? o If not, how can the results be maintained over the long term? • Is there any Investment sustainability plan for the completed project?

The evaluation questions briefly outlined above will need to be further unpacked. The Terms of Reference (ToR) should clearly detail the need for the service provider to develop an evaluation framework explaining the evaluation criteria against which the specific programme will be evaluated.

Similarly, the application of African Evaluation principles, follows five key principles that provide a guiding framework for good evaluation practice in Africa, namely;

i. The evaluation empowers Africans;

ii. The evaluation is technically robust;

iii. The evaluation is ethically sound;

iv. The evaluation is rooted in Africa, yet draws from across the world; and

v. The evaluation shows the connectedness of the world, with special attention to where humanity's footprint calls for new ideas and knowledge for change and transformation.

The table below outlines the main twenty-two African Evaluation Principles

Summary of the African Evaluation Principles 2021				
P. Powerful for Africans	**T. Technically robust**	**E. Ethically sound**	**A. Africa centric yet open**	**C. Connected with the world**
P1. Conduct an appropriate, empowering process P2. Encourage reciprocity, in- cluding mutual accountability P3. Enable learning for use-ful insights	T1. Be system- atic & analytical T2. Be transpar-ent & clear T3. Be aware of dispositions T4. Ensure a feasible evalu- ation T5. Be efficient	E1. Be sensitive to stakeholders and relationships E2. Protect the rightsof people 1.3 Safeguard diver-sity and inclusion 1.4 Address inequal-ities and power asymmetries	A1. Engage with is- sues that matter in Africa A2. Consider fra- mings and methodsfrom Africa A3. Learn and adapt from the Global South, indigenous communities, and other contexts	C1. Acknowledge interdependence and interconnect- edness C2. Foster the evaluation of sustainability in keeping with key international agreements, and with the steward- ship of nature
P4. Value and strengthen domestic capacity-ties	T6. Be culturally responsive	1.5 Be free from vested interests 1.6 Consider trade-offs		C3. Strive to con- tribute to the urgent need for sustainable and transformative change

2.1.3 Types of Evaluation

There are various types of evaluations which can be conducted in Government intervention based on the three categories of evaluations, which are Ex-ante, Midterm or Post-Ante Evaluations. The type of evaluation chosen is determined by its purpose. Evaluations form part of the policy or programme cycle. Different types of evaluations which can be conducted at each stage of the policy or programme cycle are as described in *figure 1* above.

Government Institution are required to always conduct a design evaluation, and implementation evaluation to be able to ascertain if the intervention is being well designed and implemented according to the design. Such an approach will help the MDAs, RSs and LGAs to take corrective measures on interventions that are poorly designed and implemented. **Table 3 below** lists the types of evaluations that the MDAs and LGAs should conduct for different interventions.

Table 3: Evaluation types

Evaluation Type	Objective	Timing	Possible Key Evaluation Question
Diagnostic evaluation	Focuses on understanding the current situation before the design and planning of an intervention. This evaluation is preparatory to ascertain the root causes of the problem, the potential effects, and solutions to the problem	At the beginning of the intervention – before implementation.	What is the underlying problem, the needs of the target population (including structural or systemic inequities within this target group), and the context in which intervention will function?
Design evaluation	Seeks to assess the theory of change, and logical model. This is conducted before and during the implementation of an intervention to ascertain if the design work or does not work and is likely to produce the intended outcomes.	After intervention has been designed.	How is the programme conceptualized or designed? To what extent is the programme design suitable for the achievement of the desired effect and transformative equity) within the given context?
Implementation/ Process	Evaluation focuses on the implementation process and attempts to determine how successfully the intervention is following its theory of change and logical model as designed. Provides an early warning for any problems that may occur. Allows programs to monitor how well their program plans and activities are working.	After the intervention has been implemented / started.	To what extent is the programme serving its intended participants? Who is being excluded and why? To what extent is it contributing to the required systemic changes affecting equity?

Evaluation Type	Objective	Timing	Possible Key Evaluation Question
Outcome evaluation	Focuses on measuring intervention effects in the target population (beneficiaries) by assessing the progress in the outcomes that the intervention is addressing. Tells whether the program is being effective in meeting its objectives. It is important to note the usefulness of conducting process evaluation while doing outcome evaluation. In case the outcome evaluation shows that the intervention did not produce the expected results, it may be due to intervention implementation issues.	At least a year after implementation, depending on the duration of the intervention.	What are the critical or emerging outcomes you are achieving because of the programme? What are the achieved outcomes for the marginalised groups?
Impact evaluation	Seeks to measure the changes that are tangible amongst the target beneficiaries and whether these changes are a direct result of the intervention that was implemented (attribution instead of contribution).	After 3–5 years of implementation.	To what extent has the programme changed the lives of the beneficiaries in positive ways and why? What are the benefits for the marginalised groups?
Evaluation synthesis	Seeks to systematically collect evaluation findings and generalize the results across government, particularly sectors.	At any stage but normally conducted after an intervention has been completed.	What are the cost benefits? Did all the intended beneficiaries benefit equally – including the marginalised?

Evaluation Type	Objective	Timing	Possible Key Evaluation Question
Economic evaluation	Seeks to measure the cost effectiveness of an intervention. Tells whether the program is being effective in meeting it is objectives. An economic evaluation will measure two parameters—cost and outcome (effect).	Once several evaluations have been completed	What is the evidence emerging from evaluations conducted within the sector?

2.1.4 Evaluation approaches

There are already existing approaches which can be used when conducting an evaluation. One or more evaluation approaches can be used in a single evaluation. The request for proposal (RFP) or ToR should ask the evaluators (individuals or consultancy firms) to indicate the evaluation approach to be used and its justification for utilization in an evaluation. Recently, there is a Prof. Michael Q Patton invested 'mixology' where different evaluation approaches are used. Further to that evaluation mixology incorporates theories and methods. In doing so, there should be one evaluation approach that is dominant and is used in conjunction with others. For example, an evaluator (internal or external) can adopt a utilization focused evaluation and participatory evaluation. The use of these evaluation approaches is certainly informed by evaluation questions. Table 4 presents the different evaluation approaches.

Table 4: Evaluation approaches

Approach	Key Issues	Timing
Utilisation-focused	This is based on the principle that an evaluation should be judged on its usefulness to its intended users. The intended users can be the primary users of the evaluation. These must be clearly identified and personally engaged at the beginning of the evaluation process. During the evaluation it is key to ensure that these intended users guide all other decisions during the evaluation process. This approach is aimed at maximising the utilisation of findings	From start to end of an intervention

Participatory evaluation	Participatory evaluations can be led by the intended beneficiaries of a project or programme. Intended beneficiaries are involved to some degree in decisions over the collection, analysis and use of information. It should also mainstream disability, children through targeted measures in the design, implementation, monitoring and evaluation of policies and programmes.	From start to finish of an evaluation
Theory-based Evaluation	The evaluation starts with a theory of change that shows how a project or programme should work and maps out the causal pathways between interventions and desired changes. Sometimes, the theory of change is developed before an evaluation is commissioned. Sometimes, it is developed (or adapted) as part of the evaluation[4]. It therefore seeks to test a theory.	At any time depending on availability of theory of change
Developmental evaluation	This evaluation approach can assist social innovators develop social change initiatives in complex or uncertain environments. This approach response to the context	From start to finish of an evaluation
Gender responsive	Gender-responsive evaluations assess to what extent an intervention has resulted in progress (or the lack thereof) towards intended and/or unintended results regarding gender equality. A gender-responsive evaluation should also be a process that is inclusive, participatory and respectful. This is normally used with feminist theories.	From start to finish of an evaluation
Disability response evaluation	A disability evaluation put into consideration the rights of persons with disabilities. It is key to consider the people with disability during design and implementation of interventions.	From start to finish of an intervention
Social-Equity responsive evaluation	The evaluation takes a closer look at inequalities and discriminatory structures to provide evidence about who benefits (and does not) from development policies and programmes. An absence of equity and gender focus in evaluations carries the risk of perpetuating discriminatory structures and practices	From start to finish of an evaluation

4 INTRAC (2017)

Climate Change-responsive evaluation	The evaluation takes into consideration how human and natural circumstances are directly or indirectly affecting the climate.	From start to finish of an evaluation
Empowerment evaluation	This is a stakeholder involvement approach designed to provide groups with the tools and knowledge they need to monitor and evaluate their own performance and accomplish their goals	From start to finish of an evaluation
Blue marble evaluation	The approach seeks to create a new paradigm for evaluation that aligns with the realities of an interconnected and rapidly changing world (i.e. climate change, poverty, biodiversity loss, and global health issues). By adopting this approach, evaluators can contribute to more informed and sustainable decision-making, leading to positive global impacts and a better future for all.	From start to finish of an evaluation

2.1.5 Who conducts evaluations?

In undertaking evaluations in Tanzania, internal and external evaluators should be used. For external evaluators the government procurement system which comply to the Public Procurement Act [CAP. 411 R.E. 2019], Public Finance Act [CAP.348 RE 2020]

The Local Government Finance Act. [CAP. 290 R.E.2019] must be adhered to.

2.1.6 Internal evaluators

Internal evaluators are the public Servants in the M&E Divisions, Unit or sections. They review interventions conducted in the respective MDAs, RSs and LGAs which means they must account to the Head of organization they are evaluating, or they belong to. Most of internal Evaluators are involved in the feasibility, formative or process evaluations but not limited. This has the advantage of being more familiar with the intervention and its history, as well as remaining with a project and being invested in the results, increasing the possibility of learning from the evaluation findings, and putting the recommendation to suitable use. Also, the National Coordinating Ministry's such as Prime Minister's Office –Policy, Parliament and Coordination, President's Office Regional Administrational and Local Government Authorities, Ministry of Planning and Investment, Office of the Treasurer Registrar etc. are the Internal Evaluators to evaluate different Government Interventions so as to assess performance.

2.1.7 External evaluators

These are independent evaluators who are not employed by the government of the United Republic of Tanzania or any public sector organization whose intervention is being evaluated.

These may also come from outside the country, when local independent evaluators may be limited in their skill set. For transparency and impartiality, external evaluators from outside the country may be employed, particularly in outcome and impact evaluations. External evaluators should generally have specialized knowledge that a program could benefit from, and they are often seen to be more objective. They must also bring best practices expertise from other organizations or nations for the local organization to learn.

2.1.8 Hybrid evaluations

A hybrid evaluation involves both internal and external staff working together. Hybrid evaluations is seeking to bring together the advantages of both types of evaluators. Some of the ways of organizing a hybrid evaluation shall be that:

(i) The external, and internal, evaluators shall work together in an evaluation team, which shall be led by an internal and/or external evaluation manager;

(ii) The external evaluator shall support internal staff to conduct an evaluation through facilitation and/or coaching and shall provide just-in-time technical advice;

(iii) External and internal evaluators shall, each, take responsibility for one or more components where they have particular expertise (e.g., specialist data collection or analysis, or reporting); and

(iv) The external evaluator shall undertake a quality review of the evaluation at key milestones (e.g., evaluation design, data collection instruments, evaluation report).

2.1.9 When to conduct an evaluation?

All Public Service Institutions (inclusive of the local government), all evaluations of national flagship or importance will be evaluated based on the following criteria:

(i) Budget considerations

Evaluations should be conducted if the projects fall within certain budget thresholds. Projects with a planned budget or actual expenditure of at least TZS 10 million or more must plan and undertake both a midterm (inclusive of process/implementation evaluation) and final evaluation (outcome and impact evaluations). Projects with a planned budget or actual expenditure between TZS 5 million and TZS 10 million must plan and undertake either a midterm (inclusive of process/implementation evaluation) or final evaluation (inclusive of outcome and impact evaluations). In situations where it would be strategic to conduct an evaluation of a project which has a planned budget or actual expenditure less than TZS 5 million, this should be considered, subject to approval of the Minister of the requesting ministry or accounting authority of the requesting government Entity.

(ii) Duration

Projects with the following minimum duration are eligible for evaluation. Projects with a duration of at least five years must plan for, and undertake either a mid-term or final evaluation as well as ex-post evaluation (meta and impact evaluations) whereby;

- In the mid-term of the project the mid-term evaluation should be undertaken to assess the continued relevance of an intervention and process made towards achieving planned activities; and
- In general, public Service Institutions are encouraged to conduct final/ end-line evaluation to be conducted after the completion of the project.

(iii)Scaling –up/termination of a project

All projects that have gone through the piloting stage need to be evaluated for further scaling-up (or alternatively, termination), taking into consideration what is elicited from the evaluation (either scaling or abandoning the project/programme

2.1.10 Methods of data collection during an evaluation

In conducting any of the above stated evaluation types, Tanzania public officials will need to bear in mind that there are several data collection methods which can be used. Below is a summary of key methods used in evaluations is presented below, however the application of each method is based on purpose.

Government institutions are encouraged to use mixed-methods approach (i.e., combining both quantitative and qualitative approaches). The methodology for evaluations shall be determined by the evaluation questions and the type of evaluation. Data sources utilized in each evaluation must be triangulated. Relevant monitoring data collected by the MDAs, RSs and LGAs and other publicly available data shall be utilized. Below is a summary of key methods used in evaluations:

(i) **Document review**: Review of key programme documents and reports. This also is a qualitative data collection method that involve analysing issues (evaluand) that has been already documented. These should include the key programme documents provided by the Programme team such memorandum of understandings, service level agreements, contracts, performance reports, site visit data and reports etc. Programmes and implementing partners (if any) shall give prompt access to any additional relevant key programme documentation.

(ii) **Literature review**: The evaluators shall also conduct a review of relevant non-programme literature and findings. This includes existing reports and journal articles on the evaluand. The evaluator must be cognizant of the credibility of the utilised source.

(iii) **Interviews:** Interviews with key stakeholders (e.g., programme beneficiaries, implementing agents/partners, funders): This is one of the qualitative data collection methods and it can be used to acquire more information on the targeted sample or group. (Structured, Semi-structured and Unstructured). In instances where stakeholders were previously engaged (e.g., at project initiation), the evaluation team should attempt to engage a sample of the same group (in as far as possible). Depending on the purpose, there are interviews of individuals, groups, key informants, etc. The methods applied depends on the structure of each question as shown in the box

- ✓ **Structured interview** - conducted by following a specific set of questions as in a question guide.
- ✓ **Semi-structured Interview** - conducted with a question guide, but the interviewer may not stick to the sequence of questions but adapt to the interviewees responses and asks follow-up questions for clarity.
- ✓ **Unstructured interview** - conducted in an open-ended style with questions made with a clear understanding of the objectives.

(iv) **Surveys:** A structured set of questions, usually closed ended which is used to solicit responses from respondents. It is in essence a collection of comments from a wide range of direct beneficiaries and/policy makers.

(v) **Focus Group Discussion**: Extraction of comments and ideas through a group discussion with approximately ten relevant persons on specific themes.

(vi) **Base-line study**: This is an analysis of the current situations, which is conducted prior to the implementation of policy, strategy, programme or project, with which an assessment of progress reports and comparisons are made.

(vii) **Questionnaire**: A structured set of questions, usually closed ended which is used to solicit responses from respondents. It is in essence a collection of comments from a wide range of direct beneficiaries and/policy makers.

(viii) **Case Study:** A method of developing principles of general characteristics of a specific group and clarifying these characteristics through interviews and observations as well as by accumulating similar cases.

(ix) **Mixed Method design:** In recent years, there is consensus among evaluation practitioners, regarding the importance of using several data collection methods, both quantitative and qualitative, to answer divergent questions in an evaluation. It is encouraged that evaluations in the MDAs, RSs and LGAs to make use of the mixed method using both quantitative and qualitative methods.

Table 5: Summary on advantages and disadvantages of different data collection tools

Method	Purpose	Advantages	Disadvantages
Interview	To fully understand one's impressions or experiences, or learn more about their answers to questionnaires	• Can get full range and depth of information. • Can develop relationship with respondents. • Can be flexible with client	• They are time consuming. • Can be hard to analyse and compare. • They are costly. • Interviewer biasness
Focus Group	To provide increased understanding or clarity on issues including previously obtained qualitative and quantitative data results.	• Can quickly and reliably get common impressions. • Can efficiently give significant range and depth of information in short time. • Can convey key information about programs	• Can be hard to analyse responses. • For safety and closure, it needs good facilitator. • Difficult to schedule meeting time for six to eight people
Questionnaire, Survey, and Checklist	To source information quickly or easily from people in a non-threatening way	• Can be completed anonymously. • Inexpensive to administer. • Easy to compare and analyse. • Can administer to many people. • Can get lots of data. • Many sample questionnaires already exist	• Might not get careful feedback. • Wording can bias client's responses. • Impersonal • In surveys, may need sampling and statistical expertise. • Does not get full story
Observation	To gather accurate information about how a program operates, particularly about processes	• Can view operations of a program as they are occurring. • Can adapt to events as they occur	• Can be difficult to interpret behaviours. • Observations can be difficult to categorize. • Can influence participants' behaviours. • Can be expensive

Method	Purpose	Advantages	Disadvantages
Existing Data	To gather information on the audience or the issue; to identify what previous investigators have found about the state of the knowledge, skills, behaviours, or attitudes of the intended audience with relation to the issue	• Can provide much information in little time. • Has been reviewed or seen by audience. • Makes use of already gathered information. • Helps to chart changes over time. • Provides evidence about the problem. • Involves minimum effort or interruption of audience	• Can be out of date (e.g., technology needs) • Data synthesis can be difficult. • May not address specific questions of concern. • Not flexible means to get data; data restricted to what already exists. • Statistical data may not address perceptions of the problem, or may not address causes. • Reports may be incomplete

2.1.11 Steps in data collection

Step 1: Prior to the data collection process, the evaluator develops an evaluation matrix outlining the measurement priorities, evaluation questions and the methods to be used. This should include the data collection instruments. The questions included in the evaluation matrix must be specific, measurable, and concise. They must be designed such that they qualify or disqualify solutions to the eminent problems or opportunities. Furthermore, where necessary, applications for ethical approval shall be made.

Step 2: The draft evaluation matrix, data collection instruments, and ethical approval must be reviewed by the methodological peer reviewer and evaluation steering committee.

Step 3: The evaluation manager shall consolidate all comments and inputs received on the evaluation matrix and data collection instruments. Once this is done, the evaluation manager shall share, with the evaluator, comments and inputs of the peer reviewer and evaluation steering committee.

Step 4: The evaluator shall incorporate the comments and inputs of the evaluation steering committee and peer reviewer. Once this has been done, they shall submit the revised draft evaluation matrix and data collection instruments to the evaluation steering group.

Step 5: The evaluation manager shall set up an evaluation steering committee meeting for the presentation and approval of the evaluation matrix and data collection instruments.

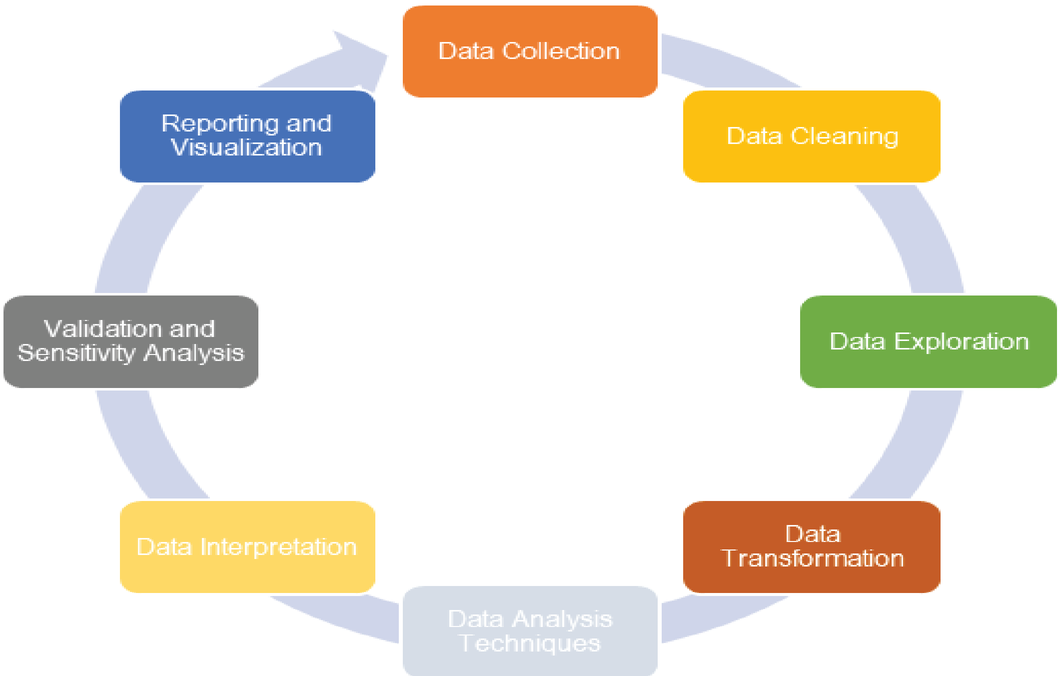
The evaluation steering committee shall approve the evaluation matrix and data collection instruments should it be satisfied that the comments and inputs have been adequately addressed. This, once ethical approval has also been sought from the requisite institution.

Step 6: Data collection process commences.

2.1.12. Data Analysis process

The evaluator shall analyse the raw data in order to source relevant information that answers the evaluation questions, and ultimately promote evidence-based policy and decision-making. Infographics may be used to present the data. However, these must be accompanied by a concise write-up explaining the contents of the infographic.

Figure 1: Data Analysis Process



Once data is collected by evaluators (MDAs, and LGAs – internal and external – consultancy firms), analysis will commence. Data analysis is a critical process that involves evaluating, cleaning, converting, and interpreting data in order to extract valuable insights and make informed decisions. Typical data analysis activities include:

- **Data Collection**: The initial stage is to collect pertinent data from multiple sources. Surveys, experiments, observations, interviews, or data acquired from databases, spreadsheets, or other sources can all be used.
- **Data Cleaning**: Following data has been acquired, it must be cleaned and pre-processed to remove any errors, inconsistencies, missing values, or unnecessary information. Cleaning data guarantees that it is of good quality and appropriate for analysis.
- **Data Exploration**: In this step, you will investigate the data to obtain a basic grasp

of its features. This can include summarizing statistics, visualizing data, and spotting patterns or trends.

- **Data Transformation:** Depending on the analysis objectives, you may need to transform the data to make it suitable for specific techniques or models.

- **Data Analysis Techniques:** Choose appropriate data analysis techniques based on the nature of the data and the research questions or objectives. For quantitative data, statistics will be generated i.e. inferential and descriptive statistics, while for qualitative data, thematic and context analysis will be performed. Different software's can be used to perform analysis, for example Statistical Package for the Social Sciences (SPSS), Microsoft excel, STATA, NVivo, ATLAS.ti and PowerBI to mention a few. Public Institutions are encouraged to leverage the existing technologies for data collection and analysis developed by E-Government Authority such as e-Dodoso survey system.

- **Data Interpretation:** After completing the analysis, analyse the results to get useful conclusions and insights from the data. It is critical to evaluate the background as well as the constraints of the analysis.

- **Validation and Sensitivity Analysis**: Validate the results by checking for consistency and running sensitivity analysis to see how changes in assumptions or input data affect the outputs.

- **Reporting and Visualization:** Present the analysis results in a straightforward and intelligible manner. Visualizations such as charts, graphs, and dashboards can aid in effectively communicating complicated data. As for the qualitative data, networks and relationship diagrams can be established using Atlas ti.

Summary of triangulation in research and evaluation

MDAs, RSs and LGAs are encouraged to triangulate[5]. It helps to generates findings that are hopefully accurate, reliable, credible, unbiased and validated. This can be done at different levels of research and evaluation. The different types of triangulations are:

- **Data triangulation**: Using data from different times, spaces, people, and sources.

- **Theory triangulation**: Using varying theoretical perspectives in research/ evaluation.

- **Investigator/evaluator triangulation**: Use of multiple researchers/evaluators in collecting or analyzing data.

- **Methodological triangulation**: Using different methods to assess the same topic for example qualitative and quantitative methods.

5 According to Better Evaluation, Triangulation is a process of validation of data through cross verification from more than two sources so as to tests the consistency of findings obtained through different instruments

CHAPTER

3 EVALUATION PLANNING

3.1 Introduction

This section outlines the steps to be taken in planning for evaluations. To promote the institutionalisation of evaluations, MDAs and LGAs are required to develop evaluation plans outlining evaluations to be undertaken within a specified period. According to the NM&E systems framework, each level of government must develop monitoring and evaluation plan.

3.1.1 Different types of Evaluation Plans

3.1.1.1 ***National Evaluation Plans (NEPs***): The PMED, in collaboration with the Monitoring, Evaluation and Reporting Technical Working Group[6], shall be responsible for the development of the National Evaluation Plans. The purpose of the National Evaluation Plan is to detail the approved strategic evaluations to be undertaken over a period of 5 years. These evaluations must be linked to the National Five-Year Development Plan, and the Tanzania Development Vision 2025. Further, the NEP shall be reviewed on an annual basis.

3.1.1.2 ***Ministries, Independent Departments and Executive Agencies Evaluation Plans (MDA-EPs):*** All MDAs shall develop their MDA-EP outlining evaluations to be conducted for a period of five years. As stipulated in the M&E systems framework and their Strategic Plans, the evaluations must be linked to the strategic plans (planning, budgeting, monitoring and reporting) of each MDAs.

For the Ministries: The M&E Unit or Section in the respective Ministry will prepare a five years Evaluation Plan and present to the Management for the approval before submitting to the PMO-PPC to develop a National Evaluation Plan.

For Executive Agencies and Independent Departments: They will prepare a five years Evaluation Plan and submit to the Parent Ministry to be compiled and submitted to the PMO-PPC to be included in the National Evaluation Plan.

For the Public Parastatals: Monitoring and Evaluation unit in the respective parastatals, with the support of Office of the Treasurer Register (OTR) shall co-ordinate the development of the five years Evaluation Plan and submit to

6 Consist of All M&E Directors from Ministries and selected Central Institutions such as NBS, NaOT, E-GA, OTR, PPRA and PO-PC

the PMO-PPC to constitute the National Evaluation Plan. Consultations with respective MDA oversight structures must be conducted as part of the MDA-EP development process.

It is important to note that all Evaluation Plan shall also be reviewed annually to align with Annual Plans and Budgets.

3.1.1.3 Regional Secretariats and Local Government Authorities Evaluation Plans (RSs and LGA-EPs): All local government authorities shall develop evaluation plans covering a given year period and submit to the President's Office – Regional Administration and Local Government Authorities who will finally submit a cumulative RSs and LGAs Evaluation Plan. The LGA-EPs shall outline evaluations to be conducted within a five-year period and reviewed on an annual basis. Further, the LGA-EPs shall be linked to the strategic plans or generated during strategic Planning process. The LGAs monitoring and evaluation unit/schedule, with the support of the PO-RALG and PMO and reporting MDAs, shall co-ordinate the development of the LGA-EP. Consultations with LGA oversight structures must be conducted as part of the LGA-EP development process.

The NEP, MDA-EP, and LGA-EP template has been attached as *Annexure A*.

3.3. Steps in Planning for NEPs, MDA-EPs, and LGA-Eps

The following section outlines the steps to be followed in evaluation planning at a national level. Important to note, these steps should also be followed by MDAs and LGAs in developing MDA-EPs and LGA-EPs. The relevant oversight authority must be consulted at the MDA and LGA level.

Step 1: Establishing / Operationalisation of the Technical Working Group

The PMED, through the existing Planning, Budgeting, Monitoring, Evaluation and Reporting Technical Working Group, will bring together the centre of government institution to facilitate the implementation of the National Evaluation System and consultations on the development of National Evaluation Plans. This includes advising on the selection and prioritisation of evaluations as well as their quality and use. As such, the roles and responsibilities of this Technical Working Group, in relation to the development of the National Evaluation Plan would be to:

(i) Develop evaluation plans to institutionalise the evaluation system;

(ii) Operationalise the set criteria for identifying, selection and prioritisation evaluations;

(iii) Recommend evaluations to form part of the National Evaluation Plan, using the set criteria;

(iv) Develop and review concept notes submitted for proposed evaluations in the National Evaluation Plan;

(v) Provide methodological advice on evaluation tools developed to support the National Evaluation System;

(vi) Act as evaluation champions in their respective institutions; and

(vii) Monitor the implementation of the National Evaluation Plan.

In developing or discussing issues relating to the National Evaluation Plan, the Technical Working Group may invite other relevant institutions, including academic institutions as well as community or beneficiary representatives.

Step 2: Issue Call for Evaluations

Once the concept note template has been finalised, a call for evaluation is issued. This is done in the form of a letter sent to MDAs and LGAs. This letter is accompanied by a concept note template (Attached as Annexure B). The MDAs, RSs and LGAs shall be requested to complete the concept note with information pertaining to the proposed evaluation for the National Evaluation Plan.

Step 3: Review and approval of Concept note by the Technical Working Group

The technical working group shall review the submitted concept notes, with the aim of ascertaining which of the proposed evaluations to include in the National Evaluation Plan. The evaluations selected as part of the evaluation plan must meet the following criteria:

1. The evaluation must be of national priority.

(i) Linked to the FYDPs - the evaluation must be critical to the development trajectory of the country.

(ii) Strategic in nature - where it is important to promote learning and replication of the intervention.

(iii) The intervention with a planned budget or actual expenditure of over TZS 500 million and above, covering a wide range of beneficiaries must plan and undertake both a midterm (inclusive of process/implementation evaluation) and final evaluation (outcome and impact evaluations).

(iv) Alignment with the planning cycle. Selection of the evaluation type should be linked to the planning cycle and the focus should be on the listed above evaluation type. Further to that is the evaluation key questions, evaluation purpose and the identified problem.

2. In addition, the proposed evaluation must meet the following criteria:

(v) Interventions with a duration of at least five years must plan for and undertake either a process/implementation or outcome evaluation or both;

(vi) Interventions that are targeting vulnerable groups i.e., women, young girls, youth and persons with disability;

(vii) Interventions that have gone through the piloting stage need to be evaluated for further scaling-up (or alternatively, termination), taking into consideration what is elicited from the evaluation (either scaling or abandoning the project/programme);

(viii) Interventions that have never been evaluated;

(ix) Interventions that have robust monitoring data;

(x) Interventions that are poorly performing;

(xi) Evaluation must be gender responsive; and

(xii) Evaluation must consider issues of climate change.

The above-mentioned criteria must be applied by all Public Service Institutions in undertaking evaluation planning. The evaluations which meet the above criteria must be included in the respective evaluation plan, i.e. NEP, MDA-EP, or RS and LGA-EP.

Step 4: Design Clinic on selected evaluations with experts

Since the utilisation focused approach to evaluation will be adopted, Evaluations must be designed with the aim of producing evidence to improve programmes, projects, and policies. It is important that evaluation concept documents are of high quality and consulted with key stakeholders affected by the evaluation. The PMO-PPC (PMED) shall host Design Clinics (Evaluation Excellence). Design clinics are used to facilitate stakeholder participation in the design and conceptualisation of evaluations to be included in the evaluation plan. It allows for a co-creative design process of proposed evaluations. The following must be completed during the design clinic:

(i) Review the Theory of Change (ToC) of the evaluation; and

(ii) Review of elements of the draft Terms of Reference (to be discussed in greater detail in the next chapter of this manual). This includes the evaluation purpose, questions, and methodology.

A thorough stakeholder analysis for each evaluation is key to the success of the design clinic. PMO-PPC (PMED) must ensure all stakeholders affected by the specific programme, policy or project are present during the design clinic are involved. This is to ensure that the evaluation is effectively design and conceptualisation. It further provides an opportunity to gain buy- in and ownership of the recommendations of the evaluation. It creates an opportunity for the stakeholders to be aware of the evaluation and form part of the evaluation process.

Figure 2: Design Clinic Process

1. Stakeholder Analysis – ascertain the key stakeholders affected by the policy, programme, or project being evaluated.
2. Drafting of the design clinic programme
3. Identify facilitators of key items on the programme
4. Invitations sent out to key participants of the design clinic
5. Facilitation of the design clinic
6. Compilation of the design clinic report

1.4 The process of developing project specific evaluation plans

The MDAs and LGAs should understand that the above evaluation plans are abstract in representing the evaluations that will be undertaken at different institutional levels. However, upon conducting each evaluation, an evaluation plan shall be developed internally or externally depending on who is conducting the evaluation. The evaluation plan will guide the MDAs and LGAs through each step of the process of evaluation. It helps them to decide what sort of information will be required by varying stakeholders. Additionally, it shall serve as an effective learning and accountability tool. Figure below is a representation of the steps of developing a project specific evaluation plan.

Figure 3: Evaluation planning process

Step One: Developing an evaluation plan.

It is vital that MDAs and LGAs present a timed and fully costed evaluation plan to the PMED, which will submit the evaluation plan to the Evaluation Reference Group (ERG) for approval. The plan should be strategic, practical, cost-effective and include evaluations of different types (project, programme, etc.) that will generate the most critical and useful information for GoT. The plan should ensure accountability and learning from implementation. When submitted to the PMED for review, all evaluation plans must be accompanied by an evaluation rationale: a brief note (maximum 300 words) explaining the justification for the evaluations included in the plan. The evaluation rationale should explain:

- How the evaluations contribute to learning and accountability, and the achievement of strategic results; and
- How the evaluations provide sufficient and balanced coverage of the programme unit's areas of engagement.

Step Two: Evaluation plan content

In deciding what to evaluate, MDAs and LGAs should first determine the purpose of proposed evaluations, as well as other factors (such as MDAs/LGAs priorities, emerging areas of engagement or potential scale-up opportunities) that may influence the relevance and use of evaluations. The evaluation plan should reflect the goals and outcomes of the Institutional, Sector or national programme and take a balanced approach, ensuring evaluation of all programmatic areas to provide the broadest accountability and learning. The contents of the evaluation plan should be checked against the following criteria:

- ✓ Planned evaluations are strategic;
- ✓ Evaluation coverage is as inclusive and balanced as possible;

- ✓ All mandatory evaluations are included;
- ✓ Inclusion of project evaluations meeting the following criteria;
- ✓ Timing, costs, resources and sequencing are realistic; and
- ✓ Influencing and constraining factors have been fully considered.

Step Three: Costing and identifying financial support for the evaluation plan

Costing of the evaluation plan is important and should be realistic, in relation to the requirements and scope of the evaluation, as well as the realities of the MDA's and LGAs budget. The PMED annual report on evaluation gives average annual costs for different types of evaluations across the country which should be used as a guide. MDAs and LGAs should estimate and indicate financial requirements and financing sources for each evaluation in the evaluation plan. When estimating the cost for an evaluation, it is important to consider the scope, depth and duration of the evaluation, as well as the composition of the planned evaluation team.

Step Four: Evaluation plan template

Table 6: Evaluation plan template

Main Heading	Sub-heading	Context
Introduction	Background and context to the activity	Briefly summarise the background to and the context for the Activity.
	Evaluation purpose	State the purpose of the evaluation.
	Evaluation scope	State the scope of the evaluation. Also include what is not in scope.
Evaluation Design		The evaluation design describes the evaluation's approach, method and tools that will be used to meet the evaluation's purpose, objectives, and key questions
	Evaluation principles underpinning this evaluation	Describe briefly in one or two paragraph(s) how the Kenya evaluation principles will be addressed.
	Information collection	For each evaluation question summarise the: type of information required to answer the question; source(s) of that information; method that will be used to gather the information

Evaluation Schedule	Timing and deliverables	Identify the key tasks to be undertaken in the evaluation, the deliverables and timing. Identify agreed progress reporting (type and frequency)
Evaluation Stakeholders	Stakeholders with interest in intervention	Include a description of the stakeholder groups in the evaluation; their interest or stake in the evaluation and whether the stakeholder group directly benefits from the Activity being evaluated (primary), or are indirectly involved with the Activity (secondary); any issues or constraints in stakeholders' participation in the evaluation (e.g. power issues, access, and confidentiality) and how this can be managed. Explain how the participation of marginalised and vulnerable communities, groups and/or beneficiaries, including women will be ensured. How the stakeholders will be involved/participate in the evaluation
Other Considerations in the Evaluation	Quality considerations	Outline how quality issues will be taken into consideration in the evaluation
	Ethical considerations	Outline how ethical issues will be taken into consideration in the evaluation. For example, do no harm; confidentiality and anonymity; gender and cultural considerations
	Limitations, risks and constraints	List potential or actual risks, limitations and constraints (e.g., around methodology, evaluation process), their likely effect on the evaluation and how they will be managed/mitigated.
	Management arrangements	Outline management arrangements for the evaluation. This includes any governance arrangements that are in place for the evaluation team

Communicating Evaluation Findings	Communications plan	Include a high-level communications plan about the evaluation for different audiences and stakeholders
	Dissemination plan	The dissemination plan will identify potential dissemination opportunities of the evaluation key findings, conclusions and recommendations to, partners, stakeholders and others

Step Five: Evaluation plan review and quality assurance process

All evaluation plans will go through an Evaluation Reference Technical Group review process. The reviewers use a checklist of requirements for the evaluation plan to verify that the criteria and requirements for the content of the plan, have been fully considered and included.

Table 7: Quality assuring the evaluation plan

Quality assurance criteria	YES	NO
Is the evaluation plan complete, i.e., noting the following? ✓ The commissioning unit ✓ Evaluation partners (only for joint evaluations) ✓ Evaluation type (programme, project, outcome, etc.) ✓ Planned evaluation completion dates ✓ Are evaluations aligned to the NSDP? ✓ Estimated budget and source of the funding		
Are all mandatory evaluations included?[7]		
Is there a brief note explaining the rationale for including the evaluations in the plan (maximum 300 words)?		
Is there inclusive and balanced coverage of the country programme content?		
Are the timing and sequencing of evaluations in the plan realistic?		
Does costing properly reflect the scope, depth and duration of each evaluation? Is it realistic?		

Step Six: Evaluation plan completion and approval

Once the evaluation plan has been finalized it will be endorsed by Monitoring and Evaluation Steering Committee in the Institutions and finally approved by the Management.

7 Refer to Question 4 addressed in chapter 2

Step Seven: Making changes to the evaluation plan

Adjustments to individual evaluations and the evaluation plan should be considered annually as part of the MDAs and LGAs stocktaking exercise. Changes that can be made with approval include:

- Extending the completion date for evaluations;
- Changing the scope and purpose of evaluations due to changes in the context (e.g., crisis settings);
- Addition of new evaluations: new projects may require new and additional evaluations that need to be included in the evaluation plan; and
- Deletion (in exceptional circumstances).

Step Eight: Monitoring compliance

Once an evaluation plan has been approved, and is entered into the GoT tracking system, the PMED, MDAs and LGAs will use the plan as a basis for monitoring compliance.

3.5. Key Stakeholders

The process of evaluation planning requires the participation of varying key stakeholders. A participatory evaluation process creates an opportunity to gain stakeholder buy-in and ownership – facilitating utilisation of the findings. The key stakeholders are given an opportunity to understand the evaluation process as well as the evaluation findings and recommendations. Ownership and buy-in allows for the effective implementation of evaluation recommendations. Below is an outline of the roles and responsibilities of key stakeholders, who must be involved in the evaluation planning process:

Table 8: Roles and responsibilities of stakeholders

Structure	Roles and responsibilities
National Monitoring, Evaluation and Reporting Technical Working Group which consist of all Head of M&E Units in the Ministries and selected Public Service Institutions	- Develop evaluation plans to institutionalise the evaluation system; - Operationalise the set criteria for identifying selection and prioritisation evaluations; - Recommend evaluations to form part of the National Evaluation Plan, using the set criteria; - Develop and review concept notes submitted for proposed evaluations in the National Evaluation Plan; - Provide methodological advice on evaluation tools developed to support the National Evaluation System; - Act as evaluation champions in their respective institutions; and - Monitor the implementation of the National Evaluation Plan.

	- Scrutinise the NEP - Mobilizes resources. - Submit National Evaluation Plan the Office of the Prime Minister and the different Parliamentary Portfolio Committees.
PMO (Performance Monitoring and Evaluation Division)	- Initiate the process for the development of the National Evaluation Plan/Agenda - Ensure the National Evaluation Plan/Agenda is included in the Monitoring, Evaluation and Reporting Technical Working Group agenda
Permanent Secretaries	- Approves the NEP - Mobilizes resources. - Submit National Evaluation Plan the Office of the Prime Minister and the different Parliamentary Portfolio Committees.
MDAs & LGAs	- Identify the evaluations that should be constituted in the NEP. - Plan and budget for evaluation identified in the NEP - Implement recommendations of an evaluation
Development Partners	- Participates in the Evaluation Technical Working Group - Co-finances the evaluations, particularly those within their scope of work and are of national interest. - Reports periodically by submitting reports to PMED especially those are sector specific and are of national interests.
President's Office – Planning Commission	- Oversees the monitoring and evaluation of National Development plans, Flagship Projects and Programmes.

1.5 Developing Evaluation Questions

The formulation of evaluation questions differs according to area of focus. Evaluation questions can generally be grouped into three categories: descriptive, normative and impact (cause-effect) questions. Many evaluations use only descriptive and normative questions. Evaluations focusing on impact ask cause-effect questions, but they typically also include some descriptive and normative questions.

(i) Descriptive questions

No	Examples of descriptive questions
1	What are the goals of the programme from the perspectives of different stakeholders?
2	What are the primary activities of the programme?
3	Where is the programme implemented?
4	Who received what services?
5	To what extent does the programme design reflect lessons learned from past similar programmes?

Usually provide a snapshot of what happened. These types of questions may also describe aspects of a process, a condition or a set of views. Descriptive questions seek to understand or describe a programme or process; they are normally straight forward (who, what, where, when, how, how many); can be used to describe inputs, activities and outputs and frequently used to gather opinions from programme clients.

(ii) Normative questions

Compare the current situation to what should be and they need some criteria with which to compare. Example of normative questions: -

(i) Are we doing what we are supposed to be doing?

(ii) Are we hitting our target?

(iii) Did we accomplish what we said we would accomplish?

(iv) Did a programme or project achieve its objectives?

(iii) Impact (Cause-Effect) Questions

This kind of questions determine what difference the intervention makes. Often referred to as attributional questions, they attempt to measure what has changed because of the project intervention although changes in the society might be caused by other interventions. Therefore, it is difficult to demonstrate that the outcomes were truly the result of that project or programme.

Examples of impact (cause and effect) questions

(i) Did the microenterprise programme reduce poverty rates in the towns in which they operated?

(ii) Did the increased tax on fuel improve air quality?

(iii) Did the increase in financial penalties for firms in violation reduce the use of under-age children in the garment industry?

(iv) As a result of the job training programme, do participants have higher paid jobs than they otherwise would have?

Evaluation questions that should be avoided

Some questions do not add value to the evaluation process and must be avoided

Type	Example	What to do
Double-Barrelled Questions	Did the project benefit you and what did you do to benefit from it?	Ask one question at a time. Do not combine questions and expect an answer
Two-in-one Questions	What are the advantages and disadvantages women participating in income-generating projects in the rural areas face in Tanzania?	Do not combine opposite positions in one question separate out the parts to make them clearer
Restrictive Questions	Do you think that female chief executives are as good as male chief executives?	This type of question eliminates the possibility that females might be better. Do not use questions that inherently eliminates some options
Leading Questions	The income-generating project has liberated women from economic dependence on their spouses. What do you think the income-generating project has done for women?	This type of question states the position or view of the interviewer. This tends to lead the respondent to the same direction. The question should be neutral
Loaded Questions	Would you favor or oppose gender mainstreaming by agreeing with the government's policy of affirmative action?	This question is loaded and emotionally charged and should be avoided. Break the question into two separate parts to make more sense out of it.

Source:

3.6. Evaluation phases

3.6.1 Evaluability assessment

Pre-evaluation of an intervention identified for evaluation entails conducting evaluability assessment based on the existing and approved evaluation plan. The assessment is conducted by the commissioning MDA or LGA to determine the feasibility of the planned evaluation, done with reference to the evaluability assessment checklist below. The evaluation assessment checklist establishes what is to be evaluated, what information is needed, what and for whom it will be used. The commissioning MDA or LGA must appoint an evaluation manager to guide, lead and coordinate the evaluation.

The OECD-DAC defines evaluability as "the extent to which an activity or project can be evaluated in a reliable and credible manner" (OECD-DAC 2010, P21). Literature on this subject shows that evaluability can be used in two ways: "in principle" evaluability is concerned with the nature of project design, including the theory of change. It looks at whether it is possible to evaluate it as it is described in its present form. Secondly, "in practice" evaluability looks at the availability of relevant data as well as systems and capacities which make the data available. As such, the purpose of evaluability assessment check elements of the programme design, implementation, and contextual issues of an evaluation. It is a key aspect of evaluation quality assurance, ensuring the cost effectiveness of evaluations. It is important that evaluability assessments are conducted at this stage of the process. This would allow for the selection of evaluations which have passed the evaluability assessment, saving state recourses.

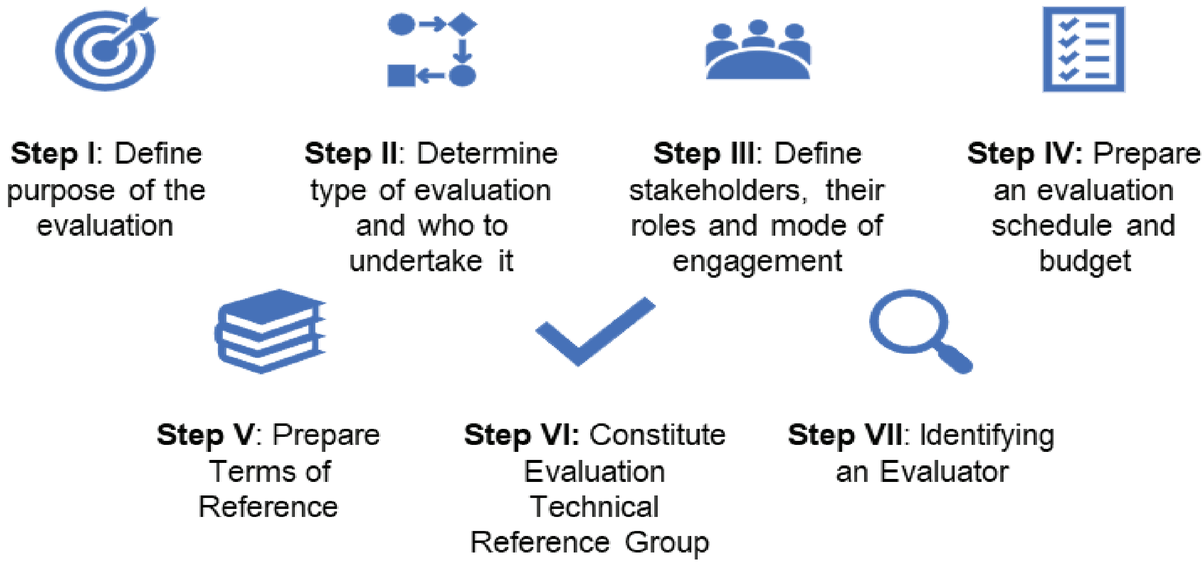
Table 9: Evaluation assessment checklist

	Evaluability Assessment Checklist	Y	N
1.	Does the subject of the evaluation have a **clearly defined theory of change**? Is there common understanding as to what initiatives will be subject to evaluation?		
2.	Is there a **well-defined results framework for the initiative(s)** that are subject to evaluation? Are goals, outcome statements, outputs, inputs and activities clearly defined? Are indicators SMART (Specific, Measurable, Assignable, Relevant and Time-bound).		
3.	Is there **sufficient data for evaluation**? This may include baseline data, data collected from monitoring against a set of targets, well-documented progress reports, field visit reports, reviews and previous evaluations.		
4.	Is the planned evaluation **still relevant**, given the evolving context? Are the purpose and scope of the evaluation clearly defined and commonly shared among stakeholders? What evaluation questions are of interest to whom? Are these questions realistic, given the project design and likely data availability and resources available for the evaluation?		
5.	**Will political, social and economic factors** allow for effective implementation and use of the evaluation as envisaged?		
6.	Are there **sufficient resources** (human and financial) allocated to the evaluation?		

3.6.2 Preparing for the evaluation

Evaluation undertaking is a cumbersome process that requires robust planning. There are various steps that are involved in the process of planning for an evaluation. However, the steps of planning are not sequential. Figure below shows the different stages of planning for an evaluation and are explained further step-by-step.

Figure 4: Preparing for evaluation



Step I: Define the Purpose of the Evaluation

The starting point and most important step in preparing for an evaluation process is to have a clear understanding of the evaluation utilisation. Ideally, the commissioning MDA or LGA should decide the evaluation purpose. The purpose may include:

- Understanding the extent to which the design and/or the implementation process of a development intervention have contributed to its success;
- Identifying the challenges or success factors;
- Identifying the conditions in which the intervention can be successfully replicated;
- Assessing whether the resources have been spent efficiently and/or effectively; and
- Assessing the intervention's impacts.

Step II: Determine the type of evaluation and who to undertake the evaluation

The commissioning MDA or LGA should determine the purpose and type of evaluation to be undertaken and who to undertake the evaluation based on the purpose/objectives.

Step III: Define the stakeholders, their roles and mode of engagement

Key stakeholders to be involved in the evaluation should be identified and their roles and modalities of engagement defined (consultation/discussion of documents, country workshop, accompanying group, comments on the draft report, etc.).

Step IV: Prepare an evaluation schedule and budget

Evaluation requires adequate financial and logistical resources. It's therefore important that the budgeting process is comprehensive and covers all costs related to the evaluation. Items to be considered during preparation of the evaluation budget include:

- Professional fees for all evaluators or thematic experts undertaking the evaluation.
- Additional and non-professional costs such as daily subsistence allowance, transport expenses.
- Translation costs for interviews, field visits, validation, and dissemination workshops.
- Any costs related to focus group meetings or data-collection meetings (venue hire, snacks, participant transport, etc.).
- Communications costs including editing, publication, and dissemination costs.
- Stakeholder, validation, or workshop costs.

Relevant program staff should be involved in the budgeting process or consulted about the budget and should be encouraged to give feedback. Once the budget is done, there is need to prepare an evaluation schedule to help the Evaluation Technical Reference Group and evaluation teams manage the evaluation.

Step V: Prepare Terms of Reference

The Terms of Reference (ToR) document defines all aspects of how a consultant, or an evaluation team will conduct an evaluation. The ToR forms the basis for a contractual arrangement and act as the main reference document during the evaluation. It will be drafted by the commissioning MDA and reviewed in consultation with ETRG to ensure high quality standards. The evaluation terms of reference must be accurate and specific. This will ensure that evaluations conducted are of high quality. In cases where the ToR is developed for a non-competitive process (for example internal), it is still advisable to provide an opportunity for discussion to clarify the expectations of all involved parties. In case of where the evaluation is to be conducted by an external evaluator, the ToR should be annexed to the contract since they are an integral part of the legal agreement with the commissioning MDA or LGA that details the extent of services, the number and quality of deliverables and the timeline for the evaluation. The terms of reference template have been attached as Annexure C.

Step VI: Constitute Evaluation Technical Reference Group (ERG)

Once the budget and the ToR have been approved, the Accounting Officer (Permanent Secretary, RAS, CEO, DED DG etc) will appoint in writing an ETRG which will be headed by the Head of Monitoring and Evaluation Unit/Section. The ETRG will comprise of technical experts in the field under which the evaluation will be conducted.

Step VII: Identifying an Evaluator

Once the ETRG has been commissioned, the commissioning MDA or LGA should initiate the recruitment process of the evaluator. The ToR forms the basis for identification/ recruitment of the evaluator as they specify the appropriate background and specific experience for evaluations. Any recruitment should be done in line with the Government procurement rules and regulations. In the event the evaluation will be conducted internally, the Accounting Officer of the commissioning MDA will constitute an evaluation team comprising of the staff from the respective MDA.

The services of an external evaluator must be procured in accordance with relevant policies and legislation pertaining to procurement. Once the decision has been made, an advertisement for potential evaluators will be issued. The advert must be accompanied by the evaluation Terms of Reference and a clear indication of the criteria for choosing the most qualified service provider. The process is outlined below:

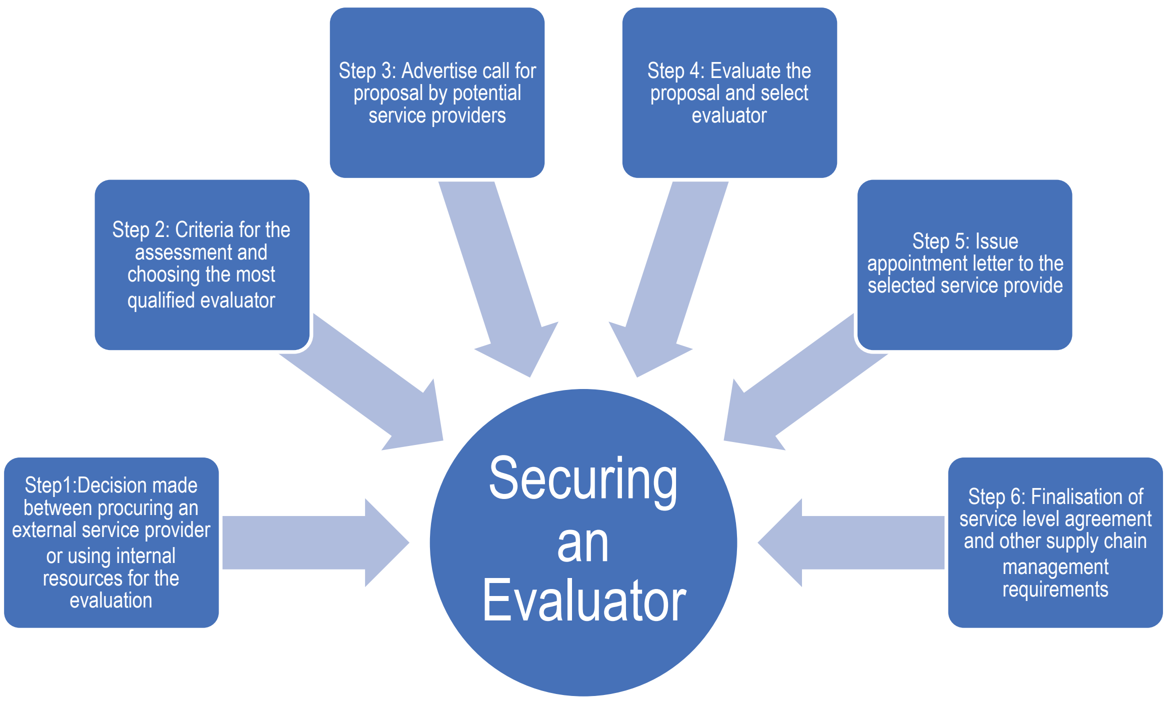


Figure 5: Process for procuring an external evaluator

The Head of Monitoring and Evaluation must ensure that this process adheres to the Government Procurement Rules and Regulations. Further, a representative from the commissioning entity and selected methodological expert who is knowledgeable of evaluating topic or intervention must form part of the selection process.

The complete terms of references, including elements developed during the design clinic, are finalised, and approved by the Evaluation Steering Committee. The Evaluation Steering Committee is comprised of key stakeholders in the sector being evaluated or affected by the evaluation. This committee is involved in each stage of the evaluation process, overseeing, and approving all deliverables of the evaluation. It is the responsibility of the Evaluation Steering Committee to ensure that the evaluation is of the highest quality and take decisions on the evaluation process. The table below outlines key structures and role players in the evaluation process.

Table 10: Composition of structures

Structure	Roles and responsibilities
Evaluation Commissioner **(Accounting Officer)**	Call for evaluations to be conducted. Secure funding for the proposed evaluations Set up an Evaluation Steering Committee Lead and oversee the evaluation
Head of Monitoring and Evaluation Unit/ Section	The evaluation manager shall be an internal staff member who is the Head of Monitoring and Evaluation Unit. They are responsible for: Oversee the strategic and operational activities of the evaluation. Monitor the implementation of the evaluation deliverables in accordance with the stipulated timelines in the work plan
Evaluation Steering Committee	Determine scope and focus of the evaluation. Review and approve project deliverables: inception report, Approve the data collection instruments and tools where applicable. Provide strategic direction and guidance of the project. Review and approve evaluation reports

Evaluation Technical Group	To manage the technical aspects of the evaluation project Review the proposals. Provide technical inputs, support, and advice. Consider recommendations from the Steering Committee. Provide continuous progress on evaluations to the Steering Committee and various stakeholders. Make recommendations to Fusion Centre as the Custodian for the evaluation report. Ensure that required deliverables are timely produced. **Work stream A: Literature review** Conduct literature search, review and prepare a write-up. Develop and apply methods; collect and analyse relevant data; write and present report. **Work stream B: Benchmarking** Develop and apply methods; collect and analyse relevant data; write and present report. **Work stream C: Evaluative workshop** Develop and apply methods for data collection through the evaluative workshop facilitate evaluative workshop. Develop and apply methods for development of Theory of Change (ToC) and review of the ToC.
Peer Reviewers	Review and provide comments on key deliverables, i.e., Inception Report, Literature Review, Theory of Change and draft reports. Review the quality and credibility of the evaluation process.

3.6.3 Key deliverables

Key deliverables expected from the evaluator and detailed in the ToR should include, at a minimum, the below:

Table 11: Key outsourced evaluation deliverables

Key deliverables	What deliverable aims to achieve
Inception Report	Detailing the understanding of the evaluation, the approach and the methodology.
Evaluation framework	Outlines the approach to the evaluation confirming the evaluation questions, the instruments to be used and the sources of information

Evaluation draft report	Draft report for review and consultation
Evaluation final report	Report including key findings and recommendations, including a condensed summary of the report (e.g., executive summary). Submission should include background information and supporting data, references, and appendices (e.g., list of interviewees).
Revised Theory of Change	The evaluation's findings and recommendations may result in the need to amend the intervention's theory of change. As such, the revised theory of change must be submitted once the evaluation has been completed.
A final power-point presentation	Outlining the evaluation framework, process, findings, and recommendations.

3.6.4. Work plan

The evaluation work plan and terms of payment must form part of the evaluation terms of reference. It projects the deliverables of the evaluation, budget allocations and the amount of time required for each deliverable. This allows for effective time and stakeholder management. The Evaluation Steering Committee can plan its meetings for the approval of deliverables, based on this work plan and payment terms. The Head of M&E must ensure that amendments to the work plan, as well as additional information reflected in the evaluation work plan template, forms part of the inception report approved by the evaluation steering committee. Furthermore, ensure that contracts have been signed with each team member. The work plan template has been attached as **Annexure D.**

CHAPTER

4 EVALUATION UNDERTAKING AND USE

4.1 Introduction

Once all the preparations of conducting an evaluation have been done. It is important to now conduct or manage the evaluation. This chapter presents ways through which the evaluation process would ensue and be managed by the commissioner of that evaluation. Additionally, the chapter presents the management response and implementation plans. This is critical to ensuring that the MDAs and LGAs can utilize the evaluative findings.

4.2 Process of managing the evaluation

Managing an evaluation is complex and strenuous. It involves managing the evaluation team as well as the stakeholders who have interest and power on intervention that is being evaluated. MDAs and LGAs should ensure that they keep the stakeholders informed during the process of initiating, implementing and post implementation of the evaluation findings. Figure 5 below presents how to manage an evaluation.

Figure 6: Managing the evaluation

Inception Meeting

- An inception meeting is held with the service provider and the evaluation steering committee
- The inception report is discussed and approved by the Steering Committee

Implementation of the Evaluation

- A literature review, data collection tools, theory of change and analytical framework, draft evaluation report are produced by the service provider
- ESC comments on each of the above and approves each deliverable
- A workshop is held with key stakeholders to validate the findings of the evaluation and discuss the recommendations

An improvement plan is produced

Approval of the Evaluation

-The ESC provides approval of the report

Peer Reviewer provides approval of the evaluation

- Quality Assurance of the evaluations
- All evaluation undergo PMED approval process

Post implementation of evaluation

-Approved evaluations are shared via the respository and other forums/ workshops

-The improvement plan is monitored on a half yearly basis for a period of two years

Step 1: Briefing on evaluation team

The technical working group shall hold a pre-inception meeting will be held with the appointed service provider. This meeting shall discuss the proposal submitted by the service provider and amendments which need to be made. To avoid delays in the evaluation process, the Head of M&E shall supply the evaluators with existing and relevant documentation, as outlined in the concept note, during the pre-inception phase. This documentation includes:

(i) Programme/ project documents;

(ii) Statistical data collected on the evaluation;

(iii) Data indicators including baselines, targets, and status;

(iv) Performance reports (and other relevant reports);

(v) Previous evaluations conducted on the evaluation or other evaluations in a related field; and

(vi) Studies conducted on the evaluation.

The availability of documents pertaining to the evaluation have a bearing on the methodology evaluation and quality of the evidence produced through the evaluation. The use of existing information, documents and research minimizes costs and is therefore highly recommended. Further, provide the required insight to answer some of the evaluation questions.

Once the pre-engagement meeting has been held and the service provider has reviewed the documentation provided, amendments can be proposed to the initial proposal submitted by the service provider. These amendments shall form part of the inception report. The inception report lists changes made to the initial proposal by the service provider and the pre-inception process that took place (i.e. meetings held with the Technical Working Group etc). The report shall include background information, a brief description of the context and objectives of the evaluation. It shall outline the conceptual framework to be adopted in conducting the evaluation. It shall include an outline of the data collection, sourcing, processing and analysis methodology, the sampling framework, and key indicators. Lastly, it shall include a timeline for the evaluation project and the draft data collection instruments.

The Head of M&E and secretariat of the ERG will set up an ERG meeting to discuss and approve the inception report. The inception report shall be approved by the ERG prior to the evaluator initiating the work. The service provider shall submit the evaluation report at least 10 days (about 1 and a half weeks) prior to the ERG meeting. This will allow members of the ERG to review the report and make the necessary recommendations or comments on it. Once the ERG comments are addressed, the Chairperson shall provide a signature approving the inception report. Depending on the work plan and payment arrangement agreed upon in the terms of reference, the service provider shall receive the first payment on the approval of the inception report. The inception report template has been attached as **Annexure E.**

Step 2: Evaluation data collection

Different data collection methods are used. These are mainly aligned with the evaluation questions and scope. This is elaborated in Section 2.7 including the advantages and disadvantages thereof.

Step 3: Drafting of the evaluation report

Once the evaluation data collection is completed, the evaluator shall produce a full evaluation report, summary report and a PowerPoint presentation. The inclusions of the summary and full report are discussed below.

Summary Report

The summary report shall be inclusive of the following sections:

1. **Introduction:** The introduction to the evaluation. This section shall comprise three paragraphs outlining the evaluation and its contextual and environmental background. It shall outline the evaluation summarized background and the methodology used in the evaluation.

2. **Key evaluation findings:** The key evaluation findings shall be included in the summary report. It is important to note that funders, politicians and institution executives may not always have sufficient time to read the entire report. Therefore, the key evaluation findings shall be included in this section. This includes evaluation findings from the literature review, case studies and overall data collection process.

3. **Conclusion**: The conclusions that the evaluation draws must be included in the summary report.

4. **Recommendations**: The main recommendations of the evaluation must be included in this section. The recommendations must be clear, concise, and inclusive of the stakeholder to be responsible for its execution.

Full Evaluation Report

The evaluation report shall be structured as follows:

1. **Title of the evaluation:** The title of the evaluation report shall include the type of evaluation and the programme, project, or policy being evaluated. For example, "The implementation evaluation of the National Youth Development Programme"

2. **Policy Summary:** The key policy issues, findings and recommendations arising from the evaluation must be included in the policy summary. This section must not be more than one page. It must be drafted keeping in mind that the policy makers, institutional administrative executives, and political executives are its primary audience.

3. **Executive Summary:** In drafting this section, the evaluator must use the structure outlined in the summary report. The main message of the report, supported by evidence and statistical information must be included in this section. It must be drafted in an influential manner as some may not read the entire evaluation but only focus on this section. The main findings and recommendations of the report shall be included in this section. The section shall be a maximum of 5 pages long.

4. **Introduction to the Evaluation**: This section shall introduce the evaluation and the context in which it took place. That is, the geographical, political, socio-economic, environmental, and historical context within which the evaluation is set. The essential context for understanding the evaluation shall be included.

5. **Purpose and objective of the evaluation**: This section must indicate the intent of the evaluation, and the evaluation. The specific evaluation objective shall be included in this section. It must indicate the issue that the evaluation is seeking to understand. Importantly, the evaluation's purpose is linked to the type being undertaken. Therefore, it will be structured based on the type of evaluation being undertaken. The reasons for the selection of the specified evaluation and its goal must be included in this section.

6. **Methodology**: The methods and techniques used to evaluate the evaluation must be outlined in this section. They shall clearly indicate the steps followed in undertaking the evaluation and ensure that the evaluation findings are of the requisite standard. The parameters and standards employed shall be clearly outlined.

7. **Evaluation Findings**: This section will stipulate the findings of the evaluation. These findings must be clear and concise. They must be linked to the evaluation questions, supported by relevant evidence sourced from the primary and secondary data collected in the evaluation. They must indicate the stakeholder or institution affected by and responsible for the finding. The evaluator may discuss the findings from the primary data collection and their meaning within the context in which the evaluation operates and compare the findings with evidence from the literature review and theoretical framework adopted.

8. **Conclusions**: This section summarizes the purpose of the evaluation, the questions asked in the evaluation and the findings thereof.

9. **Recommendations**: The recommendations of the evaluation must be clearly stated. They must be concise, indicating the responsible stakeholder or institution for the execution of the recommendation. It must be structured such that the policy and decision maker is able to understand the meaning of the recommendation and is able to utilize it in the execution of their duties.

Step 4: Approval of the evaluation report

The evaluation report must be approved by the commissioner of the evaluation. The commissioner of evaluation can be an MDA or LGA that is getting technical support from PMED as the evaluation commences. The evaluation report must be quality assured before it is disseminated and communicated to wider audience.

4.4 Use of the evaluation findings

As has been stated earlier in previous chapters, Tanzania has adopted a utilization focused approach to its monitoring and evaluation system. The findings and recommendations of evaluations conducted shall contribute to evidence-based decision and policy making. The management of MDAs and LGAs are central to making sure that evaluation findings are put to suitable use.

4.0.1. Targeted Users

The main actors on undertaking evaluations are defined by the structure of the institution. Generally, among other functions, the PMO through PMED has an overall mandate on overseeing and quality control of evaluations in the Government machinery; while M&E sections under respective ministries are mandated to provide guidance and supervise evaluations at their entities and scope. According to Patton (2008), evaluations are "done for and with specific, intended uses." The Monitoring and Evaluation Systems Framework (2014) promotes the use of evidence to "anticipate and solve management problems and respond to stakeholder demands". As such, of the evaluations include both state and non-state actors. The diagram below shows some of the users of the evaluations and results.

Target Users of Monitoring and Evaluation Evidence

4.0.2. Creating an environment for evidence use

Evidence produced through evaluations must meet the information needs and bridge information gaps. Consultations shall be undertaken to clearly articulate such information needs. It shall answer the following questions: What evidence is needed? Why is this evidence needed? When is this evidence needed? Who needs this evidence? The PDME, and the respective MDAs and LGAs M&E Coordination Units, shall ensure that the following is undertaken:

Guidance to ensure Evidence Use

Instiutitionalisation and capacitation of the M&E function	• All Public Service Institutions Executives shall ensure the existence of M&E Units • The M&E Units shall have personnel with the requisite evaluation skills • All M&E Units shall be adequately funded for operational tools and evidence generation activities
Defining objectives, information needed, and desired quality of M&E data generated	• Ascertain the National, MDA and LGA information needs • Develop a data collection and dessemination strategy and manual. This should be developed such that it can be aligned with planning and budgeting operations and requirements • Data quality should form part of the key results areas of performance appriasal and work plans
Establishment of e-platforms for sharing and accessability of M&E evidence	• PMED shall develop an e-platform to be utilised by government institutions to upload M&E evidence. This platform must be accessible by all • All government institutions shall upload M&E data and evidence in this platform
Establishing standardised reference toolkit for the mainstreaming of data demand and usage in public sector planning and budgeting	• M&E Coordination Units shall ensure availability of data in real time to encourage its use in the formmulation of programm implementation action plans
Alignment of M&E Data with overarching guidelines for public sector planning and budgeting processes	• The target users should ensure alignment with overarching planning and budgeting policies, guidelines, and legislations

8

4.4.1 Management Response

The review process includes a management reaction, which ensures that the MDAs and LGAs opinions and comprehension of the evaluation are conveyed. By assuring that the evaluation has been examined and will be considered when planning future related initiatives, it also enhances the usage of the evaluation. A management response for the

8

implementation and follow-up of the MDAs and LGAs recommendations should be included in every evaluation—internal and external—done by the MDAs and LGAs. To provide easy accessibility and to maximize the benefits of the evaluation's findings, management shall guarantee the systematic distribution, as well as storage, and management of the evaluation's output.

However, before management response, the evaluation report should have been approved by the evaluation steering committee. Once the evaluation has been completed and approved by the ERG, the evaluation manager must develop a management response. The management response stipulates the findings of the evaluation, the recommendations of the evaluation, stakeholder responsible or affected by the findings and an indication of whether the stakeholder accepts or rejects the recommendation. A management response template has been attached as **Annexure F.**

4.4.2 Who completes the management response matrix

The evaluation team should develop a matrix with the suggestions and a rough timetable or timeframe for implementation. The Evaluation Manager is then given access to the draft matrix and should consult with the appropriate MDAs and LGAs management and personnel to finalize it. To ensure that the draft matrix is included in the evaluators' deliverables, it must be agreed upon with them before the evaluation exercise commences. This must be included in the evaluation ToR.

The Head of M&E of a specific MDA or LGA should work with management of that MDA or LGA to formulate actions that clearly show how, when and by whom will the recommendations be implemented. The implementation time frame should also be unanimously agreed. The 'management response' section should include reasoning and other possibilities for ideas that have been rejected or partially accepted. Clear explanations concerning acceptance should be included here. It is imperative during the solicitation of evaluation findings feedback to take note of recommendations so that agreement is reached sooner than rejecting the recommendations at management response stage.

The follow-up activities are supposed to be linked to the formulated recommendations and should be realized in a realistic timeframe. Additionally, there should be clear roles and responsibilities of those implementing the activities. Due to the implementation of the activities recommendation, it is important that MDAs and LGAs continue to monitor and also put corrective measures.

The management response matrix is required to be completed in a more participatory. It is important that the evaluation manager involves all affected stakeholders throughout the evaluation process. This is to ensure the buy-in and ownership of the stakeholders. This ensures that the management response phase is carried out efficiently. The following are the different components that constitutes the matrix, however, evaluation team and manager, MDAs, RSs and LGAs may adapt and modify to suit their preference.

- **Evaluation recommendations**: this section is meant to provide guidance to the management in the identification of key actions required to bring about the agreed upon changes and corrections to contribute to the expected results and outcomes.

- **Responsibility of individuals or units**: these are the actors who will implement the selected actions within the suggested time frames.

- **Priority level and timeframe**: the MDAs, RSs and LGAs and the management should agree on the order of priority actions.

- **Management response**: the section indicates if the management accepts, partially accepts, or rejects the recommendations and explanation should be provided.

- **Key actions**: this section includes the planned actions to realize the recommendations. The actions could be primary and secondary actions.

- **Monitoring the implementation**: this section ensures that the evaluation findings are being utilised. The section should be populated during the course of the implementation with dates showing progress and requires regular updating.

It remains the responsibility of the MDAs, RSs and LGAs who are implementing the recommendations to provide the management response to the PMED for their record keeping.

4.5 Development of an improvement plan

One the management response process has been concluded; an improvement plan shall be developed. The improvement is linked to the findings and recommendations of the evaluation. The following template can be used for the improvement plan.

The evaluation manager shall issue an improvement plan workshop invitation to stakeholders affected by the evaluation. At the improvement plan workshop, the evaluation report shall be presented, with emphasis placed on the evaluation findings and recommendations. The participants of this workshop shall be grouped according to the findings affecting their institutions. They shall be provided with the improvement plan template to populate based on the recommendation they have been allocated to. The information to be populated includes activities to be undertaken to achieve the recommendation outlined in the evaluation report, the unit responsible for the execution of specific activities, the timeline for such execution, and indicators thereof.

The groups shall report back to plenary for comments and inputs. Once this has been done, the evaluation manager will consolidate the information into one document and submit it for approval by the evaluation custodian.

The evaluation manager shall monitor progress on the implementation plan bi-annually. Progress reports shall be provided by the enlisted stakeholder for each activity in the improvement plan. Written progress reports shall be sought from the respective stakeholders bi-annually. The progress report must be signed by the respective heads of institution and programme manager. The improvement plan template has been attached as **Annexure G.**

4.6. Dissemination of the final evaluation report

The PDME shall develop a repository for approved evaluations. The repository shall be public. The evaluation terms of reference, full evaluation report, summary evaluation report, and quality assessment report must be posted onto the repository. Further, policy briefs must be developed, posted on the repository, and shared with key stakeholders affected by the evaluation. The PDME shall further ensure that the evaluations are presented at knowledge sharing forums, organize brown bags focusing on specific evaluations, and share the evaluation on social media platforms. The evaluators shall further present their work conferences and submit to evaluation related journals.

CHAPTER

5 EVALUATION QUALITY ASSESSMENT

5.1. Introduction

For the generation of knowledge, results-based management, and accountability to program partners, evaluations of excellent quality are crucial. This manual stipulates, among other things, that MDAs, RSs and LGAs and Prime Minister's Office (PMO) must make sure that evaluations are used to guide program management and advance development goals. There is therefore increased emphasis to strengthen support the capacity of MDAs, RSs and LGAs to conduct evaluation that are stipulated in the Evaluation Plan to be able to comply with the Plan and the National Monitoring and Evaluation Policy (to be developed), improve the quality of evaluations and increase the use of evaluations by policy and decision makers.

5.2 Defining quality assessment

Quality Assessment refers to assessment of evaluation design, quality of findings and evaluative evidence and the comprehensiveness of the evaluation conclusions and recommendations. It intends to ensure adherence to the predetermined standards and criteria stipulated in evaluation documents. Quality assessment takes place throughout the evaluation process starting from evaluation of ToR and ending with final evaluation report. The quality assessment process supports the improvement of the quality of evaluative evidence including findings, coverage and scope, as well as recommendations, through the independent analysis of evaluations undertaken by identified quality assurers. Quality assessment system for all evaluation reports from MDAs, RSs and LGAs will ensure that there is similarity and consistency of evaluation reports. Additionally, this enables excellent evidence synthesis and meta-analysis in respective sector evaluations.

5.3 Purpose of quality assessment

The quality review of an evaluation report offers a critique of an evaluation's approach, the veracity of the findings and evaluative information, and the reliability of the recommendations and conclusions it draws from the methods employed during the evaluation process. The quality assurance process is made possible through using predetermined scoring system and weightings. This manual therefore encourages that the MDAs, RSs and LGAs utilize the rating system developed to rate the quality of the reports during and after the evaluation process is completed. Below are the reasons why quality assurance process is important.

(i) Improving the quality of evaluative findings to inform the contributions of interventions to development outcomes;

(ii) Strengthening accountability and transparency through impartial critique of the evaluation report by an identified quality assurer by the PMED;

(iii) Enhancing uniformity in all evaluation structure and reporting across all the evaluations conducted through the NEP, MDA-EPs RS-EPs and LGA-EPs; and

(iv) Informing the MDAs, LGAs and PMED with reliable information that helps with management response and improvement planning.

5.4 Quality assurance process

It is important to ensure that all reports are quality assured. The following is a summary of the quality assurance process.

Figure 7: Quality assurance process

5.5 Institutional arrangements

The PMED is the overall custodian of the National Evaluation System (NES) therefore will play a crucial role in providing and procuring independent quality assurers for the evaluation reports particularly those constituted in the NEP.

It is the duty of MDAs, RSs and LGAs to ensure that both internal and externally conducted evaluations are done upholding the quality standards as set by the PMED. The quality assessment process should not take more than one and half months before the process is completed including the submission of the reports on the repository. PMED is responsible for compiling a quality assurance report annually to disseminate the lessons and corrective measures that need to be instituted for evaluations to be of quality.

5.6 Components of the quality assessment

The following are the quality assessment sections to consider. However, the sections are not cut and stone but can be revised, as and when necessary.

1. Terms of reference: Does the TOR address sufficiently and clearly identify purpose, objectives, criteria, and key evaluation questions in an understandable manner? Purpose and scope are crucial to ensure the quality and credibility of evaluations. Specify the scope, including the types of evaluations it applies to and the roles and responsibilities of individuals involved. Criteria against which the quality of evaluations will be assessed is important.

2. Evaluation structure, methodology and data sources: Is the evaluation report structurally good with well identified set of objectives, criteria and methodology that is appropriate and suiting the purpose of the evaluation and scope?

3. Cross cutting issues: Does the evaluation report address the key topical cross-cutting issues such as gender and inclusion, climate change, disability and any other that stakeholders who are part of the evaluation would have identified.

4. Findings, conclusions, and recommendations: Are the evaluative findings credible and are they based on the DAC criteria as described in this document (see the DAC criteria in chap 2) and are they aligned to the proposed evaluation questions. Do the conclusions go beyond evaluative findings identify priority issues. Do the conclusions present logical judgements based on findings that are substantiated by evidence? Are the recommendations relevant to the subject and purposes of the evaluation, and are they supported by evaluation evidence?

5.7. Quality Assessment weighting

The following weighting of the quality of the evaluation report can be used. Likert scale measurement is used to measure the quality based on the personal level of agreement or disagreement with a statement. Highly satisfactory is represented by six (6), whereas highly unsatisfactory is represented by one (1). Ratings of 4, 5 and 6 considers the evaluation report of good quality.

Table 12: Quality assessment rating scale

Rubric	Description	Weighting
Highly Dissatisfied	None of the requirements were met, with severe shortcomings.	6
Dissatisfied	Most of the requirements were not met, the report has shortcomings	5
Slightly Dissatisfied	More than one requirement was not met, the evaluation report has severe short comings	4
Neutral	The requirements were partially met, with some shortcomings	3

Satisfied	All requirements were met. However, there were minor shortcomings in the evaluation report	2
Very Satisfied	All requirements were met with no shortcoming	1
Not Applicable	Not applicable	Not scored

1.8 Institutional Structure

This chapter describes the responsibilities of institutions in managing evaluation activities within Departments, units and sections of monitoring and evaluation from Local Government Authorities to the National level. In addition, the chapter defines the roles of various stakeholders in the management of evaluations from development of interventions to be evaluated; criteria to consider in the evaluation analysis; classification based on criteria; analysis and verification of evaluation to be conducted; and preparation of the Evaluation Plan.

6.1 **Evaluation Management**

Nationally, the Prime Minister Office (Policy, Parliament and Coordination) through the Department of Monitoring and Evaluation of the Government Performance are responsible of supervising the implementation of evaluations planned for a given time period. The responsibilities of the department are follows: -

(i). At the National level, the Department of Monitoring and Evaluation of Government Performance will manage the quality of evaluation at Ministries, Institutions, Government Agencies and Independent Departments levels.

(ii). At the level of sectoral Ministries and institutions, evaluation management is carried out by units or sections responsible for monitoring and evaluation. The units or sections are responsible for preparing the evaluation plan in the sector; submit it to the Ministry's Management for comments; submit to the Office of the Prime Minister for approval and implementation.

(iii). In the case of Regional Secretariats, the role of evaluation management is carried out through the Planning and Coordination Section and for Local Government Authorities this role is carried out through the Planning and Coordination Department/ M&E Units.

Responsibilities of every institutions are as detailed in the Integrated National Monitoring and Evaluation Guideline of 2024.

Prime Minister's Office
Policy, Parliament and Coordination

Government City, Mtumba
P.O. Box 980
Dodoma, Tanzania
Email: ps@pmo.go.tz
www.pmo.go.tz

Ofisi ya Waziri Mkuu