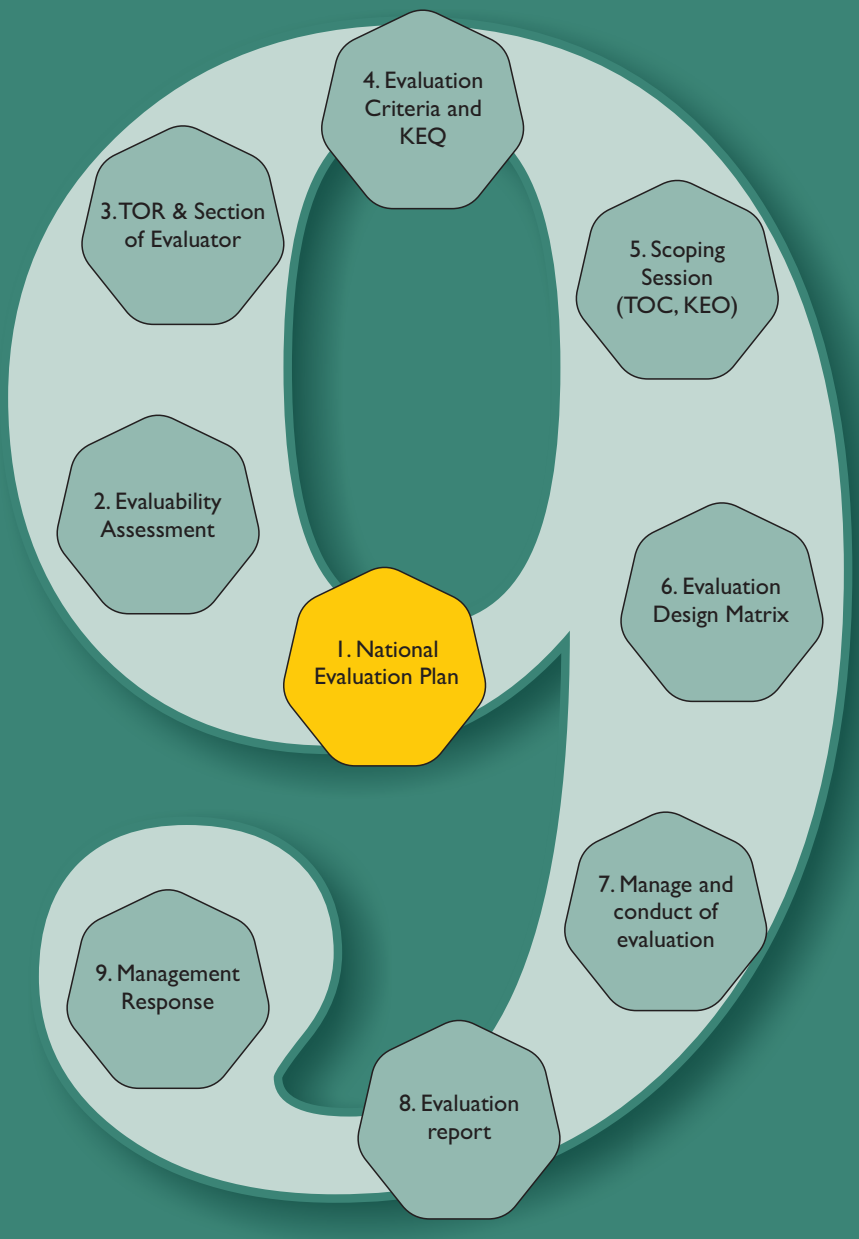
Development Evaluation in Sri Lanka

Practitioners' Guideline A Step-by-Step Approach



Development Evaluation in Sri Lanka

Practitioners' Guideline A Step-by-Step Approach

June 2020

This publication has been developed with assistance from

USAID
FROM THE AMERICAN PEOPLE

Strengthening Democratic Governance
and Accountability Project (SDGAP)
in Sri Lanka

Mr. V. Sivagnanasothy
Senior Consultant on MfDR and Development Evaluation
Former Secretary, Ministry of National Policies,
Economic Affairs, Resettlement and Rehabilitation,
Northern Province Development and Youth Affairs

About the Author

Mr. V. Sivagnanasothy served as Senior Secretary to the various Cabinet Level Ministries of Vocational Training and Skills Development, Tourism, Plantation Infrastructure Development, Prisons Reform, Resettlement and Rehabilitation, National Integration and Reconciliation, Traditional Industries and Small Enterprise Development, over the past 10 years. He last served as a Secretary to the Ministry of National Policies, Economic Affairs, Resettlement and Rehabilitation, Northern Province Development and Youth Affairs. He also parallely served as the Secretary to the Presidential Task Force on Northern and Eastern Provinces Development.

He championed the mainstreaming of MfDR in Government and Co-chaired the International Evaluation of Paris Declaration on Aid Effectiveness. He gave leadership to establish country-led national evaluation system; sensitized parliamentarians, led evaluation capacity development (ECD); established MfDR and evaluation modules in Universities and training-arms; led the formulation of national evaluation policy; led Evaluation Information System to support utility; undertook evaluations and served as Commissioner, reference group/ peer review member and user.

He served as a Vice-Chair of the Asia-Pacific Community-of-Practice on MfDR and facilitated networking and contributed articles in international conferences on MfDR and Evaluation.

He is an Economics Honors Graduate of the University of Colombo and holds MSc in Project Planning and Management of the University of Bradford, UK. He is also a Fellow Member of the Institute of Chartered Accountants of Sri Lanka and Fellow Member of the Chartered Institute of Management Accountants of U.K.
He was conferred Professional Excellence Award in 2019 by the Institute of Certified Management Accountants of Sri Lanka organized by the National Management Accountants Conference and Commonwealth Accountants Conference.

Paul King
Chief of Party
USAID'S Strengthening Democratic Governance
and Accountability Project (SDGAP)

> "When it is obvious that the goals cannot be reached, don't adjust the goals, adjust the action steps."
>
> Confucius

We live in an increasingly complex, challenging and fast-paced world. For policy makers the importance of evidence based information is critical and I am pleased to note the considerable progress that Sri Lanka has made in incorporating tools such as evaluation in their policy making process. USAID, through the SDGAP activity has been proud to partner with various government institutions in further deepening this commitment and process. Utilizing evidence impacts policy in a number of ways. First of all, it served as the foundation of recognizing an issue - the first step in the policy making process. Evidence - through the process of analysis - helps inform the design and the choice of a particular policy. It also assists in developing forecasting models to better understand how a policy impacts on both short and long term outcomes. Evidence is also critical in order to assess the expected outcome through a vigorous monitoring protocol and, of course it is at the core of evaluating a policy's impact.

Through our partnership with the institutions of governance in Sri Lanka over the past 4 years we have been able to collaboratively develop a number of tools to further the conversation and to take ideas and words and to put them into action. This has been incredibly rewarding and underscores how Sri Lankan policy makers strive to improve the lives of the fellow citizens on a daily basis. Make no mistake - the work is hard - but the commitment and drive has never ceased. As our activity closes - and new ones continue to work in partnership with Sri Lanka - we trust that the materials developed will continue to informative in the years ahead.

Foreword

Today development has become a complex process which goes beyond engineering and economic factors, covering social, environmental, gender, climate change, human rights, peace and inclusiveness. Therefore, "Development is rather a learning process in which mistakes are unavoidable despite all efforts not to make them.All involved parties must learn from past mistakes and adjust projects and plans accordingly".

In the Sustainable Development Goals (SDG) Era, the UN Resolution has strongly advocated for a country-led national evaluation system to improve performance accountability and lesson learning. It is important to understand what works, what does not work, why it works, why it does not work and in what context it works. Therefore, learning imperatives are becoming more and more important in a complex world. Further, the SDG Era demands a professional and systematic evaluation to improve accountability for results, learning and knowledge management. In this context, the UN Resolution in 2015 emphasised the need to establish and strengthen a country-led national evaluation system.

The Government of Sri Lanka, in collaboration with the private sector, civil society, academia, the non-governmental sector and Parliamentarians has demanded a strong institutionalisation and mainstreaming of development evaluation in the public sector. A Select committee of Parliament was established under the leadership of the Deputy Speaker to strengthen national evaluation capacities in the Government of Sri Lanka.

At this crucial juncture, the need to prepare a country context focused Practitioner's Guide on development evaluation with a step-by-step approach was considered a fundamental element to strengthen the mainstreaming of development evaluation. Many universities and academic institutions have dedicated post graduate programmes and special modules on development evaluation. A standardised and systematic, harmonised and professional approach to development evaluation is considered the need of the hour.

In this context, USAID came forward to support the development of the Practitioners' Guide on development evaluation.The guide was developed by Mr.V. Sivagnanasothy, Senior Evaluation Consultant and former Secretary of the Ministry, with the involvement of Ms.Vaidehi Anushaynthan, Deputy Director.

This guide will be a very useful tool to practitioners, Ministries and Departments and also to the universities and academic institutions which provide training and capacity development on development evaluation.

Acknowledgement

The Practitioner's Guide on Development Evaluation in Sri Lanka: a step-by-step approach, was developed by the Senior Evaluation Consultant Mr.V. Sivagnanasothy who served as the Senior Secretary to many Ministries in Sri Lanka. He last served as a Secretary to the Ministry of National Policies, Economic Affairs, Resettlement and Rehabilitation, Northern Province Development and Youth Affairs. Mr. V. Sivagnanasothy has worked on Development Evaluation extensively, both nationally and internationally. He served as a Co-chair of the international evaluation of the implementation of the Paris Declaration which included 11 partner country evaluations and 8 donor evaluations with the OECD initiative. He served as a Head of Evaluation at the national level and strongly advocated for country-owned, country-led national evaluation systems.

Ms.Vaidehi Anushyanthan, Deputy Director, provided extensive professional support for the development of the Practitioners' Guide on evaluation in the conceptualisation, practical case studies and methodology development. With extensive experience in development evaluation, her professional contribution to the development of this guide is especially acknowledged.

The Peer Review Team comprising of international experts, eminent persons, development partners, national experts, government officers and university academics made contributions by reviewing the document. The Practitioners' Guide on Development Evaluation in Sri Lanka is funded by the USAID supported Strengthening the Democratic Governance and Accountability Project (SDGAP). The USAID and the Strengthening the Democratic Governance and Accountability Project (SDGAP) appreciates the time and inputs provided by the peer review team during the course of the review.

Table of Contents

Acronyms

ADB	Asian Development Bank
AEA	American Evaluation Association
CSOs	Civil Society Organisations
DAC	Development Assistance Committee
EA	Evaluability Assessment
EDM	Evaluation Design Matrix
ERR	Economic Rate of Return
FOs	Farmers Organisations
ILO	International Labour Organisation
KEQs	Key Evaluation Questions
KPIs	Key Performance Indicators
LFA	Logical Framework Analysis
MfDR	Managing for Development Results
NEIS	National Evaluation Information System
NIRP	National Irrigation Rehabilitation Project
OECD	Organization for Economic Cooperation Development
O&M	Operation and Maintenance
PEU	Post Evaluation Unit
PPMS	Project Performance Management System
SDGs	Sustainable Development Goals
SDGAP	Strengthening the Democratic Governance and Accountability Project
SLEvA	Sri Lanka Evaluation Association
SLIDA	Sri Lanka Institute of Development Administration
TA	Technical Assistance
TOC	Theory of Change
TOR	Terms of Reference
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
USAID	United States Agency for International Development
VOPEs	Voluntary Organisation of Professional Evaluators
WFP	World Food Programme

CHAPTER 1

Development Evaluation: Concept, Definition and Purpose

1.1 Concept of Evaluation

A large amount of funds has been invested by developed and developing countries in development policies, programmes and projects, with a view to improving social and economic development. The experiences of several decades have revealed that the performance of many development interventions was not up to expectation. The success of a development intervention is not just measured in terms of the scale of investment, but based on the lasting impacts it has on the quality of life and standard of living of the people.

Inputs, activities and outputs are elements of an intervention and not measures of success. The success of an intervention is measured by the outcomes and impacts. Sound planning and design, effective implementation and adequate post implementation arrangements influence the success of development interventions. It is important to understand what interventions work well, what does not work well and why. Development evaluation helps to establish **"accountability for results"** and improves **"learning"** from success and failures.

According to the American Evaluation Association (AEA), the term evaluation encompasses the process of assessing the strong and weak points of programmes, policies, personnel, processes, products and organisations in order to improve their effectiveness. Evaluation is action-oriented. It is done in order to determine the value or impact of a policy, programme, practice, intervention or service and formulate recommendations to promote a change for betterment.

Michael Screven defines evaluations as "processes to determine the merit, worth and value of things; Evaluator must be able to make appraising judgements rather than to measure things".

Development is rather a learning process in which mistakes are unavoidable despite all efforts not to make them. All involved parties must learn from past mistakes and adjust projects and plans accordingly. There must be a **learning** focus rather than a **punitive** one. So, penalise people not because they make mistakes, but if they don't learn from the mistakes.

1.2 Definition of Evaluation

Evaluation is a critical analysis of the achievements and results of a project, programme, policy or institution. The OECD/DAC define evaluation as:

> Evaluation is a critical analysis of the achievements and results of a project, programme, policy or institution.

"An assessment, as systematic as possible, of the on-going or completed projects…it's design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An Evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision–making process" (DAC/OECD).

OECD/DAC definitions have been adapted by all partner countries including Sri Lanka. Sri Lanka, like many other countries, uses the term "appraisal" for prospective assessment conducted prior to approval of the intervention. However, some countries use the term as "ex-ante evaluation" for appraisal. Therefore, evaluation is an assessment of the on-going and completed development intervention and does not include interventions that are still in the planning stage.

Definition of Evaluation

- An evaluation is an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using appropriate criteria such as relevance, effectiveness, efficiency, impact and sustainability. An evaluation should provide credible, useful, evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision-making processes of organisations and stakeholders (UNEG).

- The purposes of evaluation are to promote accountability and learning. Evaluation aims to understand why—and to what extent—intended and unintended results were achieved and to analyse the implications of the results. Evaluation can inform planning, programming, budgeting, implementation and reporting and can contribute to evidence-based policymaking, development effectiveness and organisational effectiveness (UNEG).

1.3 Purpose of Evaluation

Evaluation provides a learning opportunity to find out what intervention is working well, what intervention is not working well and why. It helps to improve the quality and effectiveness of development interventions. The following are the purposes for which an evaluation is conducted

- Accountability for Results: Evaluation examines whether the public spending has made a change for betterment.
- Learning from successes and failures: Evaluation helps continuous improvement in policies and programmes.
- Knowledge Management: Evaluation helps to increase knowledge about what works, what does not work and why.
- Improves decision-making: Evaluation helps decision-makers to terminate or scale-down unsuccessful interventions and to scale-up successful interventions, and enables modification of programmes and interventions to make them more results oriented.

Focus of Evaluation:
What intervention works, what does not work and why? In what context does it work?

1.4 Types of Evaluations

1.4.1 Formative and Summative Evaluations

Evaluation can be Formative or Summative. Formative evaluation is an evaluation undertaken during the implementation of a project or programme with a view to effect improvements to the interventions. Formative evaluation is generally any evaluation that takes place before or during a project implementation, with the aim of improving designs and performance.

On-going evaluation is known as mid-term or interim or formative evaluation and is undertaken during implementation. Programmes, projects and interventions may be running into serious problems during implementation and may require a fresh look at the original objectives and strategies used to address the problem. Similarly, projects and programmes may encounter risks and require modifications, amendments and improvements during implementation. For this purpose, the Formative evaluation or on-going evaluation contributes in a big way. Formative evaluation therefore helps to focus on programme improvement.

Formative Evaluations help programme improvement whereas Summative Evaluations help to judge the programme's merit or worth.

On the other hand, summative evaluations are end-of-programme evaluations and ex-post evaluations which help to judge the programme's overall worth or merit. Summative evaluations look at the impact of the interventions on the target group and finds out what the project has achieved. Summative evaluation helps to make decisions about the programme for future adoption. Michael Screven cited "When the cook tastes the soup that's formative evaluation; when the guest tastes it, that's summative evaluation". Undue emphasis on summative evaluation can be unfortunate because without formative evaluation, the development process is incomplete and inefficient. **Formative evaluation** looks at **what needs to be improved**, whereas **Summative evaluation** looks at **what results occurred.**

Evaluations are categorised into 4 categories namely

1. Ex-ante evaluation
2. On-going evaluation
3. End-of-programme evaluation
4. Ex-post evaluation and,
5. Impact evaluation.

1.4.2 Ex-Ante Evaluation:

This is an evaluation conducted after the project has been prepared but prior to its approval, to examine the proposed interventions from a funder or sponsor's perspective to find out whether the project is really viable or feasible and is a bankable project. It is more widely known as "appraisal". It is an independent re-assessment of the viability and designs of the intervention submitted by the project proponent and it appraises the proposed project from the perspective of the funder. It includes re-assessment of the viability or feasibility in terms of marketing, technical, financial, economic, social, environmental, institutional, policy and political, legal, gender and sustainability.

Ex-ante evaluation or appraisal will examine whether the "Theory of Change" (TOC) is sound and likely to achieve outcomes. It also examines whether the assumptions are realistic and in the case of "killer assumptions", whether they have been internalised into the projects and programmes, to address the risks. The intervention gets approved after the clearance through ex-ante evaluation or appraisal.

1.4.3 On-going Evaluation

This is an evaluation conducted during the implementation of the project/programme to see whether the strategies, components and activities are appropriate and it helps mid-course corrections during implementation. On-going evaluation helps to improve the programmes based on the initial implementation experience and provides an opportunity for the planners and implementers to modify, amend and adjust the programme strategies to ensure the achievement of results. It is undertaken during implementation and it is called a mid-term or interim evaluation.

1.4.4 End-of-Programme Evaluation

This Evaluation is undertaken at the completion of the project, programme or intervention. The assets created under the project and their handing-over to the implementing agencies and the use of such assets are examined in the end-of-programme evaluation. Though it is somewhat similar to a stock-taking exercise and assesses the assets created through the programme, it also examines the possible arrangements in place to use the assets efficiently and effectively and to ensure the readiness with the operation and maintenance aspects.

1.4.5 Ex-Post Evaluation

This is an Evaluation generally undertaken two to three years after the completion of the programmes or projects, to see whether the intervention has achieved the project purpose, expected outcomes and the potential impacts, including sustainability. The ex-post evaluation also examines both intended and unintended outcomes and results. It fulfills the obligations on "performance accountability" and also helps in learning lessons for the future. The lessons learnt from the ex-post evaluations are fed into the planning process and are especially used in designing new interventions.

1.4.6 Impact Evaluation

Impact evaluation attempts to assess the long-term impacts of the programme or project. It examines whether the programme or interventions have contributed effectively to achieve the envisaged project objectives. In recent years, the development community has become aware that projects that appeared to have been implemented effectively at project completion often seem to have run into severe difficulties, later on. Such projects have not reached the targeted poor people but rather, have benefited only rich people. Hence, it has become crucial to undertake impact evaluations to verify the achievement of long-term results and the impacts of the projects and programmes, and to examine whether the long-term results have reached the targeted and intended beneficiaries. The impact assessment covers both intended and unintended impacts.

1.5 Other Types of Evaluations

1.5.1 Joint Evaluation

An evaluation in which different donor agencies and/or partners participate (OECD/DAC). In a truly collaborative joint evaluation, decisions will be made jointly or collaboratively by donors and partner countries, on evaluation planning, management, scope of the work, the team selection, methodology and reporting arrangements. Joint evaluations with partner country governments help to increase ownership for evaluation, facilitate lesson learning and help to build national evaluation capacities. Joint evaluations of different donors can also help to overcome attribution problems in assessing the programme.

1.5.2 Meta Evaluation

This term is used for evaluations designed to aggregate findings from a series of evaluations. It is also widely used to denote the "Evaluation of an evaluation", to judge its quality and/or assess the performance

of evaluators. Meta evaluation has many definitions. Some use it to describe aggregating information from several individual evaluations. Others define it as a systematic tool for quality control of evaluation studies.

Joint Evaluation helps to improve country ownership whereas Meta Evaluation helps to evaluate the evaluations to improve quality.

Meta evaluation is all about evaluating the evaluations. The quality of evaluations can be safeguarded by processes that will evaluate evaluations. Meta evaluation can prevent risk connected with unsatisfactory professional level evaluators and ensure corrections to improve the credibility and quality of the evaluation. It also ensures that evaluation principles, standards, norms and ethics are followed. It can prevent the loss of objectivity and independence that could emerge from pursuing interests that are other than professional. It can also prevent low validity of the findings in the context of an evaluation.

Michael Quinn Patton suggests that questions to focus meta evaluation should include (a) was the evaluation well done; (b) Is it worth using; (c) did the evaluation meet the professional standards and principles? It examines the quality of the evaluation processes and products. Meta evaluation is all about judging the evaluation against a set of ideas and features of what constitutes a good evaluation.

1.5.3 Participatory Evaluation

This is an Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation. Participatory evaluation is a partnership approach to evaluation in which stakeholders actively engage in developing the evaluation and all phases of its implementation. Participation occurs throughout the evaluation process including

- Identification of relevant Key Evaluation Questions (KEQs)
- Planning the evaluation design
- Selecting appropriate measures and data collection methods
- Gathering and analysing the data
- Reaching consensus about findings, conclusions and recommendations
- Disseminating results and preparing an action plan to improve programme performance.

The process also seeks to honour the perceptions, voices, preferences and decisions of the least powerful beneficiaries. These are the voices from the field. It identifies the locally relevant evaluation questions; build capacity of evaluation design and implementation; agreement on the findings and how they will use the findings.

1.5.4 Theory Based Evaluation

An Evaluation that tracks the anticipated sequence of linkages from inputs and activities to outcomes and impacts. It examines the mental model and visualised map that demonstrates the logic of the intervention and the project strategies adapted to achieve the objectives. It assesses whether are we doing the right thing and establishes whether the intended inputs, planned activities and expected results are strongly linked and aligned. It reviews whether all possible risks are taken into consideration in the design of the projects and programmes.

Theory Based Evaluations assess the causal linkages between intended inputs, planned activities and expected results and analyse the underlying assumptions and contextual factors that influence the results chain.

Over the last 20 years, theory-based approaches have increasingly moved into the mainstream of thinking and practice about how interventions are designed, measured and evaluated. It has helped to overcome the limitations of experimental evaluation designs. Evaluators are faced with two broad challenges

1. Measuring the expected results of an intervention
2. Attributing those results to the activities of the intervention

Although experimental and quasi-experimental designs are powerful to address the two challenges described above, there are several shortcomings. From a practicality point of view, it may be difficult to develop counterfactual for ethical reasons. More resources and time may be required. Experimental designs may not be able to answer questions such as "What activities of the intervention contributed to the expected results? For underperforming interventions, was the underlying theory wrong or was it a problem of implementation? Was the logic model wrong or was it due to aspirations, risks and context that supported or hindered the theory of change and its outcomes? The theory of change examines not only the causal link between outputs and outcomes but also the underlying assumptions, risks, contexts and contextual factors which contribute or hinder the achievement of outcomes.

The Theory of Change explains the relationship between the intervention and the expected results through the logic model and also examines the underlying assumptions, risks, external factors, political, social, environmental factors that may influence or hinder the achievement of expected results.

A key strength in the theory-based evaluation is the analysis of context, risks, assumptions and external factors which can enhance or hinder the workings of the intervention. The theory-based approach can provide information about which aspects of the intervention worked well and in what circumstances, what other factors affected the interventions and what might be holding it back from working better. In the absence of a TOC, the evaluator should develop a TOC. It involves three steps:

1. Developing a logic model with a clear results chain and explicit causal link.
2. Identifying the assumptions and risks underlying the Theory of Change.
3. Identifying other contextual factors associated with the results chain

Theory based evaluations allow the evaluators to arrive at findings on why interventions are working or not. Conclusions are drawn on cause-effect elements of an intervention.

1.5.5 Thematic Evaluation:

The term denotes the evaluation of a selected theme of the development intervention. For example, the evaluation of poverty, environment, gender and human rights which are of a cross cutting nature in several programmatic interventions, are evaluated as thematic evaluations.

1.5.6 Internal and External Evaluations

Internal evaluations are evaluations conducted by the programme staff, whereas external evaluations are conducted by outsiders. External evaluation can bring greater credibility, objectivity, impartiality and independence. Internal evaluations are conducted by the internal staff who know the programme history, the decision-making style of the organisation and they communicate technical results more clearly to the management. In Malaysia, self-evaluations are more regular and operational, whereas in other countries such as USA, South Africa and even among donors, the external evaluations with independence and objectivity are more popular and common.

1.6 Approaches to Evaluation

There are three approaches to evaluation. They are:

— Goal Based Evaluation
— Goal Free Evaluation
— Utilisation Focused Evaluation

1.6.1 Goal Based Evaluation

Goal based evaluation determines the extent to which the objectives of a program are actually achieved. A specification and measurement of programme goals remains central to most evaluations; what the programme or project is proposing to do and consequently how we determine whether the programme is doing what it says it is going to do. Many evaluation scholars feel that the project or programme goal is necessary if the intervention strategies and activities are relevant and connected to the outcomes through the "logic model" and the "theory of change". It is a general notion that without knowing the project goal, how do we evaluate? Therefore, evaluators continue to use the goal as their common approach for evaluation.

Setting goals, checking for consistency and evaluating the achievement of goals and feedback are the key steps in goal-based evaluations. Inadequate understanding and knowledge of social problems makes the defining of goals difficult.

Goal-based evaluations may ignore unintended outcomes, as the focus of the evaluation is centered on the intended objectives and outcomes. Goal-oriented methods limit the influence of evaluator bias by using programme identified goals as criteria for measurement. Setting goals is a necessary and sufficient condition for goal achievement, when strategies and activities are aligned with the goals.

1.6.2 Goal Free Evaluation

Goal-free evaluations are assessments in which the evaluator conducts the evaluation without reference to the stated or pre-determined goals and objectives. Michael Screven stated that pre-determined goals narrow the focus of the evaluation. Goal-free evaluations focus on actual outcomes rather than intended program outcomes. Goals act like blinders, causing the evaluator to miss important outcomes. Michel Screven states that making evaluators responsible for the use of evaluation results is also wrong. Asking them to participate in decision-making injures their independence. They will become tailored to what the client wants.

Goals are broad statements of a programme purpose or expected outcomes usually not specific enough to be measured and often concerned with long-term rather than short-term expectations. Goal-free-evaluations attempt to observe and measure all actual outcomes, effects and impacts both intended or unintended.

According to Michael Screven, the goal-free evaluation is designed to evaluate what a programme actually does, rather than what it is trying to do. The procedure followed by a goal-free evaluation involved avoiding direct contact with programme staff and avoiding discussions about programme goals. However, goal-free evaluation lacks a usable methodology.

1.6.3 Utilisation-Focused Evaluation

The utilization-focused evaluation was developed by Michael Quinn Patton and is an approach based on the principle that an evaluation should be judged on its usefulness to its intended users. Such evaluations begin with the premise that evaluations should be judged by their utility and actual use. The term **"use"**

implies how real people in the real world apply evaluation findings, experiences and lessons learnt from evaluation process.

Michael Patton stressed that the first step in evaluation is to identify decision-makers and information users, to determine what information is needed. Evaluation adds value only if its results are used. Evaluators must guide decision-makers, policy-makers and managers to do the right thing.

Utility is the prime criteria for judging the worth or merit of the evaluation regardless of its technical, practical and ethical merit.

The job of the evaluator is to facilitate the evaluation in such a way as to best meet the needs of the primary users and engage them throughout the evaluation process. The primary intended users of the evaluation must be clearly identified and fully engaged at the beginning of the evaluation process, to enhance ownership. Michael Quinn Patton argues that the intended users are likely to use the evaluation if they have ownership in the evaluation process.

The following checklist provides the key points to ensure the effective implementation of utilization-focused evaluation

a) Engage primary intended users in the evaluation.
b) The first step in utility-focused evaluation is to identify the intended users and determine what information is needed.
c) Formulate priority Key Evaluation Questions (KEQs) with the close engagement of primary users.
d) Ensure credible methods are used in the evaluation, in consultation with the primary users.
e) Stimulate the use of the evaluation findings.
f) Ensure the effective use of evaluation findings and lessons in the decision-making process.
g) Reporting and wider dissemination should influence decision-making.

The utilisation-focused evaluation looks into how an evaluation can be used effectively and how to avoid evaluation reports that gather dust on book shelves, unread and unused. Utilisation-focused evaluations are seen as a problem-solving approach that call for creative adaptations to change and be active, reactive and adaptive to the real-world situation.

1.6.4 Developmental Evaluation

Developmental evaluation involves scanning the environment, raising deep questions on the logic model, the theory of change, underlying assumptions, risks, external factors, contextual and extraneous factors, data and information systems. It investigates on a continuous basis, to facilitate real-time learning and feedback, and to ensure continuous improvement by incorporating innovation, radical programme re-design, adaptation, evolving and development of new projects or modification and amendments to existing projects, to ensure the achievement of outcomes.

According to Michael Quinn Patton, developmental evaluation is an evaluation approach that can assist social innovators to develop a social change initiative in a complex and uncertain environment.

Michael Quinn Patton considers developmental evaluation as something closer to the role of R&D in a private sector, which ensures continuous improvement in the project development process. It facilitates real-time feedback to programme staff; supports a continuous development loop; is responsive to the context and supports innovation and radical programme re-design and policy reforms with rapid real-time interaction and learning with a view to adapt to dynamic and complex environments.

Developmental Evaluation provides continuous feedback and learning to adapt in a dynamic and complex environment.

1.7 How does Evaluation differ from Monitoring?

The United Nations ACC Task Force defines monitoring as "the continuous or periodic review and surveillance (overseeing) by management at every level of the hierarchy of the implementation of an intervention, to ensure that input deliveries, work schedules, targeted outputs and other required actions are proceeding according to plan". Monitoring provides feedback on the progress of the implementation of the interventions and helps to address the problems faced during implementation.

Monitoring may be nothing more than a simple recording of intervention activities, outputs and results against the implementation plan and budget, whereas evaluation is primarily an in-depth analysis probing deeper into the planning and design, implementation and post implementation aspects including the project/programme settings, to assess the achievement of project objectives.

Monitoring provides feedback on implementation, whereas evaluation examines deeply the design, implementation and results of the programme.

The monitoring exercise examines the financial progress, physical progress, benefits and results and more importantly, focuses on implementation issues that hinder progress and addresses these through a troubleshooting arrangement to ensure that projects are implemented as planned, without time and cost overruns.

Figure 1: Monitoring and Evaluation

Monitoring	Monitoring
Concerned with efficient delivery of inputs, activities and outputs.	Concerned with the attainment of project/ programme objectives.
Tracks the progress against the implementation plan	Examines relevance, coherence, efficiency, effectiveness, impact and sustainability and uses multiple methods such as Cost-Benefit Analysis, Logical Framework Analysis (LFA), theory of change (TOC) and Evaluation Design Matrix covering quantitative, qualitative and mixed methods
Focus: Inputs, activities and outputs	Focus: Outcomes, impacts and sustainability
Measures: "Efficiency"	Measures: "Effectiveness"
Conducted by the Implementing Agency	Conducted by an Independent Consultant/ Research Institution
Source of Data: Project Staff	Source of Data: Project Beneficiaries.
Purpose: Mid-course correction during implementation.	Purpose: Learning device for future planning
More frequent and routine	Less frequent and periodic

CHAPTER 2

National Evaluation Policy Framework

2.1 Institutionalisation of a Post Evaluation System in the Ministry of Policy Planning and Implementation in the Government of Sri Lanka

As a National Focal Point of Evaluation, the then Ministry of Policy Planning and Implementation collaborated with the Asian Development Bank (ADB) and the UNDP and obtained a technical assistance (TA) programme in 1991 to strengthen post-evaluation capacity in Sri Lanka. The following interventions were completed under the TA.

a) Sensitisation of Parliamentarians and policy-makers to create demand for evaluation.

b) Conducted on-the-job evaluation training programmes of selected priority development projects and programmes to institutionalise the evaluation methodology and practices in the Government Ministries and Departments.

c) Establishment of an evaluation system and guidelines to institutionalise evaluation in the public sector, with appropriate methodology and procedures for post evaluation.

d) Introduced a curriculum on ex-post evaluation at the Sri Lanka Institute for Development Administration (SLIDA)—the key Government training arm of the public sector, to ensure continuous training and capacity-building in post evaluation.

e) Establishment of a voluntary organisation for Professional Evaluators (VOPEs) Society known as the Sri Lanka Evaluation Association (SLEvA), to serve as a non-Governmental forum and as a community of practice on development evaluation.

f) Establishment of a National Evaluation Information System (EIS)—a computerised evaluation information system to store, retrieve and disseminate evaluation findings, lessons and recommendations.

The Post Evaluation Unit (PEU) conducted over 20 post-evaluations and disseminated the findings widely, to ensure feedback to the planning, budgeting, programme management, auditing and policy-making stakeholders, to effectively use the findings, lessons and recommendations.

The TA project evaluated by the ADB was rated as successful. Based on the success of the post-evaluation functions in the Government of Sri Lanka, the ADB offered a second TA Programme to strengthen the "project performance evaluation capacity" of the then Ministry of Policy Planning and Implementation.

As a follow-up initiative, the Project Performance Management System (PPMS) was introduced and supported to identify and establish key performance indicators (KPIs) at the project and programme level.

2.2 The National Evaluation Policy

The draft National Evaluation Policy was developed in early 2003 as a further step towards the institutionalisation of the evaluation system in the Government of Sri Lanka. It was published in Dr. Ray Rist's publication "Ten Steps to RBM", as an emerging good practice.
https://openknowledge.worldbank.org/handle/10986/14926

However, in 2018, the Cabinet of Ministers granted approval to the National Evaluation Policy in Sri Lanka.

2.3 The National Evaluation Bill and the proposed National Evaluation Commission

The Ministry of National Policies and Economic Affairs gave leadership in formulating a National Evaluation Bill, in close collaboration with the Ministry of Finance and the Sri Lankan Parliamentarian Forum on Evaluation, and other stakeholders. A twelve member select committee was appointed by the Parliament of the Democratic Socialist Republic of Sri Lanka, headed by Hon. Deputy Speaker. An officials' committee supported the "Parliamentary Select Committee" and the officials' committee submitted a draft bill to establish a National Evaluation Commission (NEC) to strengthen evaluation in the public sector in Sri Lanka. The draft National Evaluation Bill was submitted to the Parliament by the Parliamentary Select Committee. The role of the National Evaluation Commission is: to select strategic projects and programmes for evaluation using a set of criteria; develop a three-year rolling plan for evaluation; set evaluation norms and standards, criteria, principles and ethics and ensure that independent, credible and high quality evaluations are conducted and that the findings and lessons are effectively used by Parliamentarians, policy-makers, planners, budgeting authorities and programme managers. The National Evaluation Bill was placed in the 8th Parliament of Democratic Socialist Republic of Sri Lanka (fourth Session).

2.4 Advocacy and sensitisation of policy-makers and senior Government Officials on Evaluation

The following universities and academic institutions have institutionalised the evaluation module into their regular curriculum.

i. **The Postgraduate Institute of Management of the University of Sri Jayawardenapura** introduced and institutionalised Evaluation into the Masters of Public Management Programmes in two modules, namely: (a) Managing for Development Results and (b) Development Evaluation, as part of their regular programme curriculum.

ii. The Master's Degree in Financial Economics of the **University of Colombo** was upgraded with the inclusion of Evaluation as an important component in the "Project Planning and Appraisal Module".

iii. **The Faculty of Graduate Studies of the University of Colombo** introduced and institutionalised an advanced evaluation module into the Masters programme on Regional Development Planning (MRDP).

iv. **The Sri Lanka Institute of Development Administration (SLIDA)** has introduced both Development Evaluation and MfDR through their regular training programmes for public officers.

v. **The Centre for Evaluation of the University of Sri Jayawardenepura** has a special postgraduate programme on Development Evaluation, which provides in-depth knowledge on development evaluation.

Similarly, the Sri Lanka Evaluation Association (SLEvA)—one of the VOPEs, conducts many ad-hoc knowledge-sharing programmes and international biennial conferences to build evaluation capacities. The Sri Lankan Parliamentarian Forum on Evaluation also undertakes many sensitisation programmes for Parliamentarians, as well as at provincial and district level.

Managing for Development Results (MfDR) and the development of KPIs are seen as an integral part of the evaluability process. Therefore, approval was granted by the Cabinet of Ministers to institutionalise MfDR in national, sectoral and ministerial plans and programmes.

In line with the approval granted by the Cabinet of Ministers to strengthen the MfDR in 2019, a circular (CS/SLAS/CPA) was issued by the Ministry of Public Administration, Home Affairs, Provincial Councils and Local Government on 17th Dec 2019, requesting SLIDA and Training Arms to build skills and competencies in the public service, in MfDR and KPIs.

2.5 Evaluation Principles

The Sri Lankan Evaluation Policy is based on five inter-related key principles, within which evaluations are undertaken.

2.5.1 Independence

- The evaluation function should be structurally independent from the operational management and decision-making function in the organization, so that it is free from undue influence, more objective and impartial and free of bias.
- Evaluators must not be directly involved in policy-setting, design, implementation or management of the subject of the evaluation, in order to avoid a conflict of interest. OECD/DAC defines Independence as "an evaluation carried out by entities and persons free of control of those responsible for design and implementation of development interventions".
- Evaluation should be objective, impartial, free from undue influence and should have full autonomy in carrying out investigations and reporting findings to the appropriate levels of decision-making.
- The evaluator should have full access to all relevant information required for evaluation and should be able to work freely, without any interference.
- Management must not impose any restrictions on the scope, content, comments and recommendations of the evaluation report.
- Evaluators must have freedom from political influence and organisational pressures and should not have any vested interest.
- The evaluator should have full authority to submit reports directly to the appropriate levels of decision-making.

Independence may be seen as Organisational Independence, behavioural independence and process or methodological independence.

1. **Organisational Independence**
   Ensure that the Evaluation function is located independently from other management functions: The following aspects are vital for organisational independence.
   - Who commissions the evaluation?
   - Is the evaluation function located in the Executive Branch or Legislative Branch of the State? In USA, the evaluation is located in the Government Accountability Office (GAO), which falls under the legislative Branch and reports to Congress. Therefore, there are checks and balances. In South

Africa, the Department of Planning, Monitoring and Evaluation (DPME) comes under the President's Office and the Executive Branch has responsibility for M&E.

2. **Behavioral Independence**
   Evaluators should possess professional competencies, skills, expertise, capacities and ethical behaviour, to ensure behavioural independence.

3. **Process or Methodological Independence**
   - Independence can be enhanced by managing the evaluation process.
   - Involving stakeholders, reference groups and peer reviewers, compliance with evaluation standards, ethics, norms, codes of conduct, common guidelines and quality assurance and using rigorous methodology will help to ensure methodological independence. Multiple methods are encouraged to overcome methodological weaknesses.

Issues and Challenges

- Independence vs Influencing: There is a tradeoff between independence on one hand and integration, learning and feedback on the other. Overly independent outsider-driven evaluations may sometime have difficulty in integrating the lessons and findings into the planning, policy making and programme management process.
- Accountability vs Learning: If accountability is the purpose, then independence is vital. But if learning is the objective, then over-emphasising independence may adversely affect the learning.
- Independence will help to increase the credibility and trustworthiness of the evaluation and will also help to increase the use of evaluation.

2.5.2 Credibility

Credibility is grounded in independence, impartiality and in the use of rigorous methodology. Credibility involves transparent, inclusive and robust methods with a quality assurance system.

Credibility requires that evaluations should report successes as well as failures. Beneficiaries should take an active part and engage in the evaluation. Credibility depends on the expertise and independence of the evaluators, the degree of transparency of the process and the quality of the evaluation outputs. Evaluations should be designed, planned and implemented in compliance with international quality standards. Explicity in the use of methodologies for data collection, analysis and interpretation will improve transparency and credibility.

Evaluations should be conducted in an ethical way with a professional approach. The Evaluator should have access to all required information. Data integrity, reliability and authenticity and the use of quantitative and qualitative data and information are all important. The data should "speak" and opinion-based conclusions should not be drawn.

Credibility is examined based on the validity of the source of information used, limitations regarding the representativeness of samples and techniques used for data collection. The following features are important to improve credibility.

- Impartiality, freedom from bias, rigorous evidence and transparency are key elements to improve credibility.
- Stakeholder participation, inclusiveness, ownership and listening to the voices of the weakest are all important.
- Democratic participation improves credibility and inclusiveness. Thus, gender balance, equity-based,

top-down and bottom-up stakeholder participation with the empowerment of citizens and the engagement ofVOPEs or CSOs enhances credibility.

- Institutional independence and the details of Commissioners, systems and processes, evaluability assessment,TOR, the selection of Evaluators, codes, guidelines, norms, instruments and mechanisms all help to ensure credibility.
- Institutional capacity, specially the capacity of Evaluators and Commissioners, their core competencies, expertise and skills in evaluation, training and professionalisation (code of conduct, capabilities, accreditation) and Peer Reviews will all influence credibility.
- Credible Evidence and Rigorous Methodology: Credible evidence, rigorous methodology with quantitative and qualitative data supported by Randomized Control Trials (RCTs) and Mixed Methods (Methodological pluralism), multiple data sources and triangulation help to ensure credibility.
- Communication and Reporting: Accessibility of the Evaluation Report, the report being publicly available and the inclusion of an acknowledgement of disagreements may also support credibility.
- Who evaluates the Evaluator: Meta Evaluations should be in place to ensure compliance with standardised procedures and processes, to ensure quality and to improve credibility.
- Evaluation findings are anchored to data and less perception based: Findings flow logically from the analysis of data, showing a clear line of evidence to support conclusions. Robustness of data, KPIs and Sound Monitoring Systems help to improve credibility.
- Incorporation of stakeholder comments: The final report should reflect the relevant stakeholder comments and acknowledge any substantive disagreements.This enhances credibility.

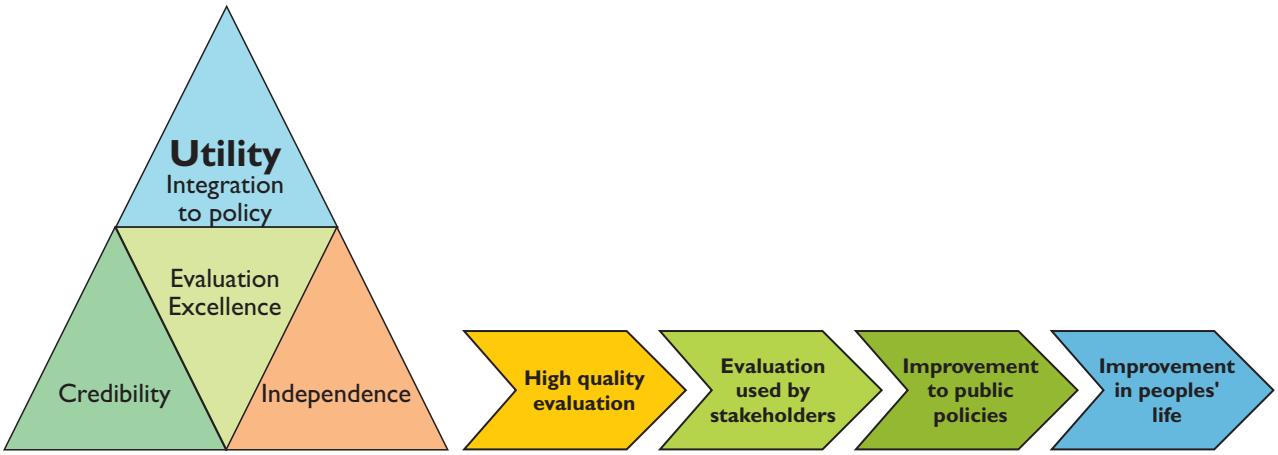
2.5.3 Utility

In conducting and commissioning evaluations, there should be a clear intention to use the resulting analysis, conclusions, or recommendations to inform decisions and actions.This utility is ensured through the following measures:

- Impact on Decision Making: For Evaluation to have an impact on decision making, the findings must be perceived as relevant and useful and be presented clearly.
- Interests of different stakeholders: The Evaluation should reflect the different interests and needs of different stakeholders, to enhance utility.
- Demand for Evaluation: There should be incentives to motivate potential users to demand evaluation. Is civil society demanding such information for both accountability and learning? The reports should be made publicly available.
- Linking to the planning, budgeting and policy process: Are the results of the evaluation taken in by the planning and budgeting authorities in the following round of the planning cycle? Is the timing of the evaluation appropriate? Are the **evaluation recommendations** strategic in nature to support decision making and enables changes in policies.
- Strategic Dissemination of Findings: The findings should be disseminated to the stakeholders, with the executive summaries.
- Evaluation Repository: Evaluation data bases and information systems should be accessed online and used by the wider group of stakeholders.
- Involvement of users in Evaluation Planning at an early stage: Getting the potential users actively involved in the formulation of evaluation questions enables effective utilisation.
- Making Evaluations Mandatory through the National Evaluation Policy: Evaluation should be an integral part of the development policy cycle. National policy should make evaluation mandatory and build an evaluation culture in the country.
- Sensitising the importance and use of Evaluation to policy-makers: The awareness and sensitisation of policy-makers to the importance of evaluation is important to create a demand for evaluation. The independence, credibility and competence of the evaluators are vital to enhance trust, which helps to enhance the use of the evaluation findings.

- National Evaluation Capacity: There should be adequate capacity to not only manage and conduct evaluations, but also to use the evaluation findings.
- Evaluation findings are timely and strategic: The evaluation findings should be strategic, actionable and reach the policy-makers in time, so that they could be included in the annual planning and budgeting cycle. Evaluation should be viewed as an "agent of change".

Figure 2: Triple Elements for Excellence in Evaluation



2.5.4 Impartiality

The key elements of impartiality are objectivity, professional integrity and an absence of bias. Impartiality should exist in all stages of the evaluation process, including planning the evaluation, selection of the evaluation team, conducting the evaluation, formulating findings and recommendations and reporting. Impartiality implies the independence of the evaluator.

2.5.5 Transparency

Transparency is vital in establishing trust and building confidence. Stakeholder engagement, ownership and increased public accountability improve transparency. The evaluation report should be accessible to all, in order to ensure transparency.

2.5.6 Human Rights and Gender Equality

The universally recognised values and principles of human rights and gender equality need to be integrated into all stages of an evaluation (UNEG). Evaluators and Evaluation Managers should ensure that these values are respected, addressed and promoted and also that the principle of "no one left behind" is operational.

2.5.7 Ethics

Evaluation must be conducted with the highest standards of integrity and respect for the beliefs, manners and customs of the social and cultural environment. Human rights, gender equality and "do no harm" principles are essential ingredients of ethics. Evaluators must respect the rights of the institutions and individuals while providing information in confidence. The Evaluator must ensure that sensitive data is protected and that consent has been obtained for the use of private information provided.

2.5.8 Professionalism

The evaluation should be conducted with professionalism and integrity. Professionalism includes credibility, the use of competencies, knowledge and skills, education, training and experience and adherence to ethics, norms and standards.

2.5.9 National Evaluation Capacities

Building national evaluation skills, competencies, capacities at country level is important to improve performance accountability and learning. The UN General Assembly Resolution emphasised the need to strengthen and support the national evaluation capacities of member states. This is important to promote countries to achieve the 2030 development agenda of the Sustainable Development Goals (SDGs).

Nine Steps to Development Evaluation

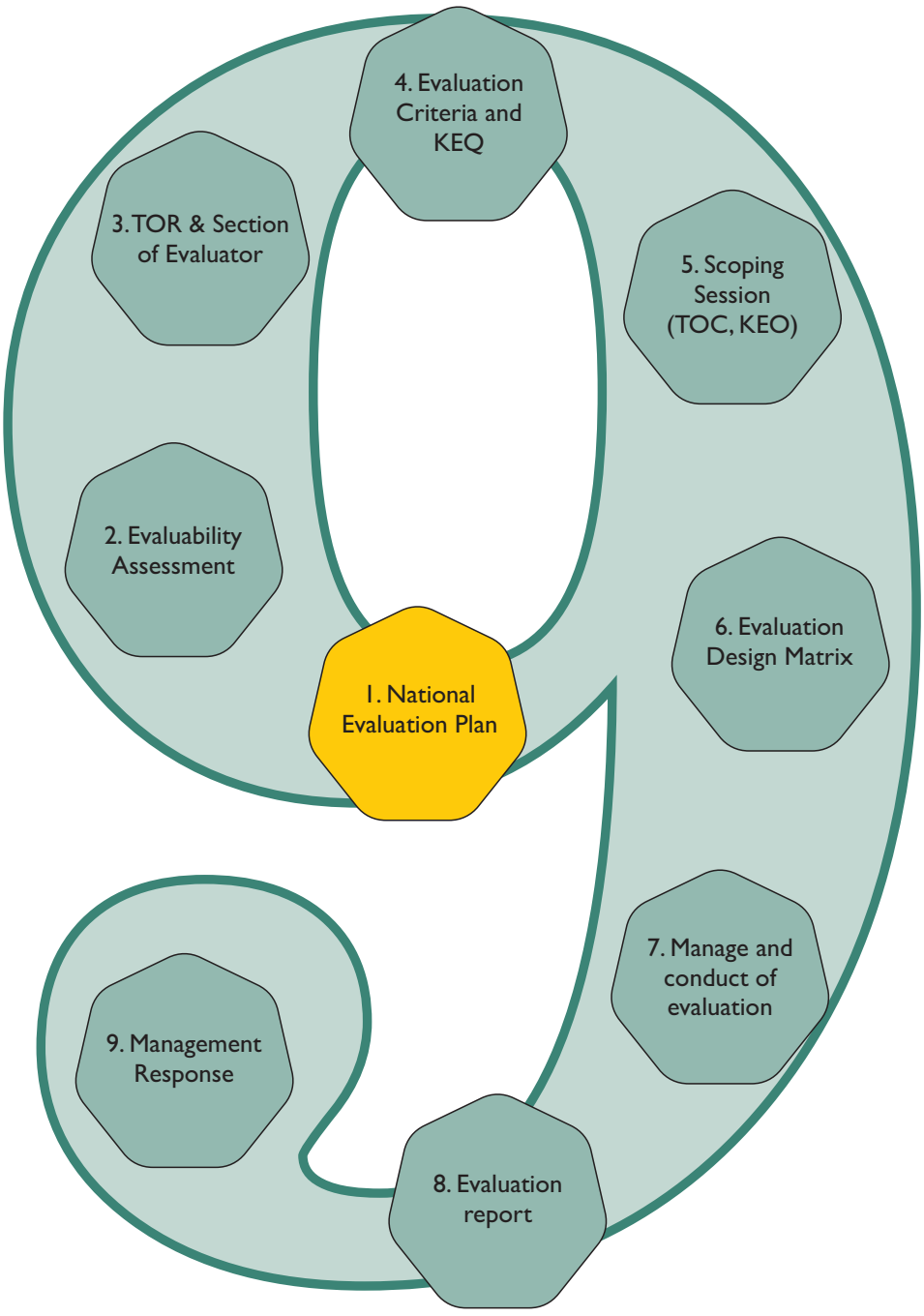


Figure 3: Nine Steps to Development Evaluation in a Sri Lankan Context

Key Steps	Key Activities
National Evaluation Plan: National Evaluation Plan (NEP), Ministry Evaluation Plan (MEP) / Provincial Evaluation Plan	Line Ministries should submit the Concept Note with a rationale and criteria for the selection of projects for evaluation. The NEP summarises the intervention to be evaluated as a national priority over a three-year period (rolling plan) with a rationale, required approval, funding, timeline, uses and users.
Evaluability Assessment	The Readiness for Evaluation of the intervention is assessed through Evaluability Assessment. It examines readiness through plausibility, feasibility, credibility and utility. **Plausible**: Examines the soundness of the Theory of Change and intervention Logic that articulates the links between activities and outcomes. **Feasible**: Examines the availability of sound data and the Information Management System and "measuring what is worth measuring". **Credible**: Examines the availability of authentic, reliable and credible data and information. **Utility**: Examines the use and usefulness of the evaluation findings and lessons in decision-amaking.
TOR and Selection of Evaluator: Procurement of Service Provider	Call for submission of evaluation proposals from potential service providers with a clear TOR and selecting the evaluators by considering the Competency Framework and the quality of the proposals.
Evaluation Criteria and Key Evaluation Questions (KEQs)	Framing Key Evaluation Questions (KEQs) under OECD/DAC Evaluation criteria, namely: Relevance, Coherence, Efficiency, Effectiveness, Impact and Sustainability (R/C/E/E/I/S) and using the Theory of Change (TOC) to design KEQs.
Scoping Session on Theory of Change (TOC), agreement by stakeholders on KEQs, based on criteria and methodology.	Scoping workshop to obtain consensus and agreement on the Theory of Change (TOC), and confirming the KEQs and methodology including design options, and ensuring their relevance and usefulness to the primary stakeholders.
Evaluation Design Matrix	Provides the Road Map for evaluation and covers Evaluation Criteria, KEQs, KPIs, data and information system, data collection methodology, design options and source of information including analysis for conclusions.
Management and Conduct of the Evaluation	A Steering Committee, Management Group and Reference Group with Peer Review to oversee the execution of the EDM. The service provider undertakes a Field Mission to collect data and information.
Evaluation Report	Report under 1/5/25 structure One Page – High Level Executive Summary 5-page - Executive summary 25-page report
Management Response, Dissemination, Feedback and Improvement Plan	Sharing Evaluation Results; communicate findings; dissemination and feedback; management response and an improvement plan to ensure the integration of the evaluation findings into policy.

CHAPTER 4

Step 1: National Evaluation Plan

The first step to implement the National Evaluation Policy Framework is the formulation of a National Evaluation Plan. The National Evaluation Plan is a three-year rolling plan which sets out the prioritised strategic key interventions, programmes and projects to be evaluated over a period of three years.

Unlike monitoring, evaluation cannot cover all the programmes and projects, as it is an expensive exercise. A call to request a programme/project evaluation should be sent by the Ministry of Finance, Economic and Policy Development or the proposed National Evaluation Commission, requesting all Line Ministries to identify priority projects for evaluation in the medium term. Finalisation of the National Evaluation Plan is based on the agreed evaluation selection criteria, as set out in para 4.5 below.

A National Evaluation Plan should be recommended by a high-powered committee and endorsed by the Cabinet of Ministers. At the country level, there are three categories of evaluation plans.

1. The National Evaluation Plan
2. The Ministry Level Evaluation Plan
3. The Provincial Evaluation Plan

4.1 The National Evaluation Plan

The National Evaluation Plan will include the evaluation proposals submitted by the Line Ministries to the Ministry of Finance, Economy and Policy Development or to the proposed National Evaluation Commission. These are subjected to a screening process using a selection criteria. Selected high priority projects are included in the 3-year rolling National Evaluation Plan for evaluation.

4.2 The Ministry–level Evaluation Plan

The Ministry-level Evaluation Plan will set out the priority interventions identified at Ministry-level for evaluation, either for learning or accountability purposes. The Ministry-level Evaluation plan should be a three-year medium-term plan approved by the Secretary of the Ministry, and may include strategic and priority programmes and projects for evaluation which are not captured in the NEP. The selected projects implemented by Departments and Statutory Boards identified for evaluation may be captured in the Ministry Evaluation Plan. It is necessary to ensure the required budgetary allocation is provided to undertake the evaluations set out in the Ministry Evaluation Plan. Such an evaluation plan will cover on-going evaluations, end of programme evaluations, ex-post and impact evaluations and it is considered useful by the Ministries for learning and possible replication and for future decisions for upscaling , down-scaling or for termination, depending on their outcomes.

4.3 Provincial Evaluation Plan

The Provinces may also develop a Provincial Evaluation Plan for prioritised strategic projects and programmes to be evaluated by the Provincial Council. A call to request a programme/ project evaluation should be sent by

the Chief Secretary of the Provincial Council requesting Provincial Line Ministries to identify priority projects for evaluation in the medium term. The finalisation of a Provincial Evaluation Plan is based on the "agreed evaluation selection criteria". The Chief Secretary of the Provincial Council, in consultation with the Governor of the Province, may approve the Provincial Evaluation Plan and undertake evaluation of the selected programmes.

4.4 Purpose of the NEP

The purpose of the evaluation plan at national, ministerial and provincial level is to list out the evaluation of projects approved by the relevant authorities as priority evaluations to be undertaken during the medium term of three years. These evaluations will serve the purpose of accountability for results as well as learning. Such an evaluation plan will cover on-going, end of programme evaluations, ex-post and impact evaluations. The outline of the National Evaluation Plan is given below.

Figure 3: Nine Steps to Development Evaluation in a Sri Lankan Context

Name of the Ministry	Intervention to be Evaluated	Rationale	Type of Evaluation	Timeframe	Budget	Users

4.5 Criteria for Selection of Projects for Evaluation

It is important to select the most strategic and priority projects for evaluation from a learning and accountability perspective. The following criteria are used in the selection of programmes, projects and interventions for evaluation.

- Projects of a strategic nature or public interest.
- Projects that are likely to be replicated.
- Projects of an innovative nature or unusual nature, where feedback is sought.
- Projects that may be running into problems and which may require a decision to terminate or re-adjust.
- Projects which may throw light on new or upcoming policy initiatives.
- Projects of high value which may be categorised as mega projects.

By applying the above criteria, it is necessary to prioritise projects for evaluation from the list of projects submitted for evaluations. Out of the prioritised projects for evaluation, some projects may not be appropriate for evaluation as they may not satisfy the evaluability criteria.

4.6 Conditions under which evaluation studies are inappropriate

- One Time Programme: No potential for continuation and not replicable. There is insufficient impact to warrant evaluation.
- Absence of commitment to use the Evaluation Results: Projects for which administrators are unwilling **to make changes** due to personal or political reasons.
- Strong Pre-Conceived Notion: Client or sponsor is not open to contradictions as there is a strong pre-conceived notion about the project.

- Absence of valid and dependable information.
- Evaluation is premature for the stage of the programme—e.g. premature summative evaluations or impact evaluations. The timing of the evaluation may be inappropriate.
- Consensus cannot be achieved among major stakeholders on the programme model and desired evaluation plan.
- The proposed evaluation is not feasible, due to inadequate financial and human resources to undertake the evaluation.

The outline of the evaluation plan set out in figure 4 above, once approved with the list of agreed projects, serves as a basis for evaluation. The next step is to undertake an evaluability assessment, which is a readiness assessment forming part of the evaluation planning process, to examine whether a feasible, credible, reliable and useful evaluation can be undertaken in terms of the likely benefits of evaluation, the uses and its cost. The next chapter will discuss the importance, benefits and methodology of undertaking the evaluability assessment.

5
CHAPTER

Step 2: Evaluability Assessment

The Evaluability Assessment (EA) examines "the extent to which an activity or project or programme can be evaluated in a reliable and credible fashion, to help to decide whether an evaluation is worthwhile in terms of its likely benefits, consequences and costs" (OECD/DAC). It helps to assess whether the programme or intervention is ready for evaluation.

An Evaluability Assessment (EA) calls for an early review of a proposed project or programme, in order to ascertain whether its objectives are adequately defined and verifiable. It is a second opinion of the programme, based on the strength of its design and logic. It is worthwhile to undertake an evaluability assessment at an early stage when something can be done to remedy weaknesses.

EA is a tool that can inform decision-makers whether the programme is suitable for an evaluation and also which type of evaluation would be most feasible, useful and credible.

5.1 Rationale

An Evaluability Assessment ensures the use of the Evaluation by facilitating communication between the Evaluators and Stakeholders. Outcome evaluations can be resource intensive and should be undertaken if the results of the evaluation are going to be used in decision-making. In the early 1970s, Wholey and his colleagues at the US Department of Health, Education and Welfare saw that the proliferation of programme evaluation in the 1960s had not resulted in an increase in the use of programme evaluation for decision-making. This led to a stockpile of evaluation reports that were not put to use. Many potential users of evaluation were unhappy with such studies, believing that they often failed to provide useful information. Wholey and his colleagues developed Evaluability Assessment as a tool to remedy this problem. They saw this as a means to facilitate communication between evaluators and stakeholders, to determine whether the programme was evaluable and to focus the evaluation study to ensure effective usage.

Evaluability assessments are methods of examining readiness for evaluation. They examine whether the goals, objectives and outputs are clearly defined; whether the "Theory of Change" of the intervention is well designed and clearly articulated; whether the key performance indicators are in place; whether the information needs of the intended users are clearly defined and agreed upon; whether the required data are accessible and whether the evaluation results will be used by the intended users. Evaluability assessment is a set of procedures for planning evaluation in order to maximise utility. It also assesses the timing of the evaluation to improve the utility of the evaluation. The objective of the EA is to assess whether a programme is evaluable.

5.2 Purpose of Evaluability Assessment

The overall purpose of an Evaluability Assessment is to decide whether the evaluation is worthwhile, in terms of its likely benefits, consequences and costs. EA was originally developed in USA in the 1970s as a way to reduce the waste associated with evaluating social programmes that were so poorly designed or implemented that no impact could be realistically expected. EA is a way of addressing the increasing number of poor-quality

evaluations. Poor evaluations are attributed to a lack of logic in the design, vague objectives, undefined indicators, data gaps, the absence of an adequate M&E framework etc. In instances of such weak design and poor M&E framework, the evaluability assessment re-works the intervention logic and recommends improvement to the M&E Framework. EA examines the nature of the project design in principle, and data availability and the data management system to carry-out evaluation in practice. It also examines the usefulness of the evaluation.

The EA identifies who needs what information, when, for whom and for what uses.

The determinants of evaluability are:

1. Clarity of the intended programme model or theory, programme goals, objectives and examine whether the performance criteria/indicators are defined and agreed.
2. Availability of data on performance.
3. A clear definition of the intended uses and users of the evaluation information.

This examines the theory, data and use.

The developers of the Evaluability Assessment believed that many evaluations had failed because they were not actionable in reality. The barriers to implement evaluation findings are attributed to non-actionable recommendations; failure to link programme activities to goals and unwillingness to make programme changes based on evaluation information. Wholey wanted to develop a way to remedy these problems. Evaluability Assessment was developed to help programmes to meet the four criteria deemed necessary for meaningful evaluation.

i. Programme goals and priority information including performance criteria are well defined and agreed upon.
ii. The programme logic model or theory of change is likely to achieve programme goals and objectives.
iii. Relevant performance data can be obtained at reasonable cost.
iv. The intended users of the evaluation have agreed on how they would use the information (Wholey).

These steps are achieved in close collaboration with the intended users. The role of the evaluator is to facilitate the discussions and to listen and learn about the programme and its information needs and expectations, from the stakeholders. Reviews of documentation, site visits and stakeholder consultations are important to undertake an evaluability assessment.

The evaluator will agree with the stakeholders and users on the goals and objectives of the programme , theory of change, underlying assumptions, contextual factors and risks, performance indicators, timing of evaluation, alternative evaluation plans, evaluation questions that need to be answered – questions of strategic interest and corresponding information requirement, who is going to use what information to make what decisions, methodology to be used, and data to be collected, quality and appropriateness of the existing measurements, project monitoring framework, identification of potential evidence gap and plausibility of the programme to achieve the goals. It will ensure value for money for the evaluation.

The two major components of the Evaluability Assessment are: (i) the Evaluability Review which aims to strengthen the project design and (ii) the M & E Appraisal which aims to strengthen the M&E Framework.

5.3 Evaluability Assessment Methodology

Evaluability Assessment can revisit and improve the logic model or theory of change or Programme model to strengthen the design of the interventions. Evaluability Assessment uses less resources than Outcome Evaluation. It can take between one to 20 months to complete.

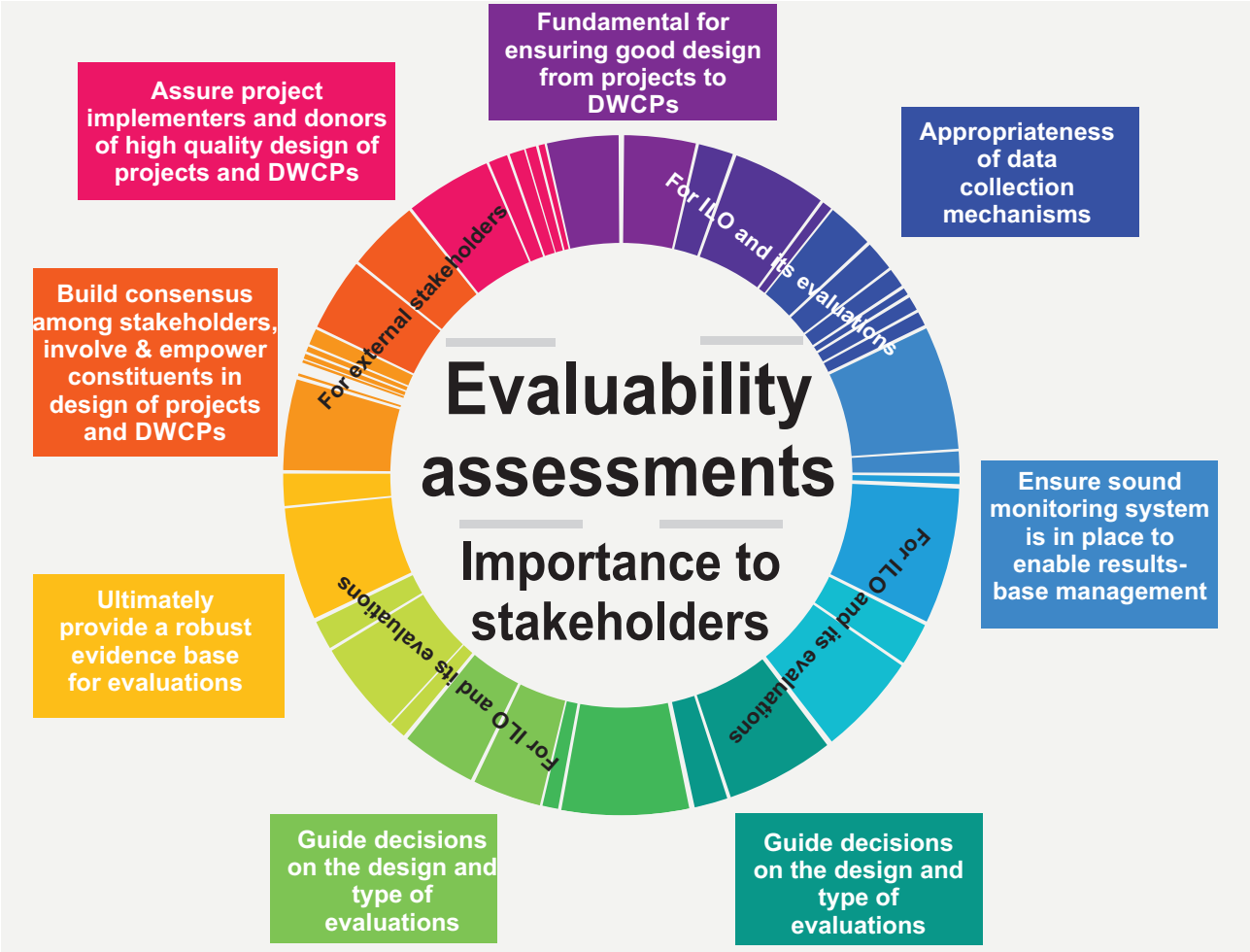
Evaluability Assessment helps to improve the theory of change, the results framework (horizontal and vertical) including Key Performance Indicators (KPIs), key assumptions and risks and results statements. It can also help to improve the design of monitoring systems, monitoring and evaluation components and the use of the evaluation. The findings and recommendations of an Evaluability Assessment should help to improve:

- The Theory of Change
- Logical Framework Analysis (Horizontal and Vertical)
- Results Statements, KPIs, baselines and targets
- Design of Monitoring System
- Key Evaluation Questions (KEQs)
- Evaluation Design Matrix, Approach and Methodology

Strengthening the project design, (evaluability review) and monitoring and evaluation framework (M&E Appraisal) are the two important aspects of Evaluability Assessment.

One option is the logic model which often visually demonstrates the relationship between the activities and intended outcomes of the programme. A "Theory of Change" goes further by describing the context and underlying assumptions of the interventions. A Programme model may also include a table of performance indicators.

Figure 5: Importance of Evaluability Assessments to Stakeholders
(Source: ILO 2019, Guidance Note 16 on Evaluability)



The Evaluability Assessment answers the following questions:

1. **Plausibility**: Plausibility examines whether the quality of the design and the activities and outputs of the project will reasonably lead to the intended outcomes. Is it plausible to expect outcomes and impacts? Does the quality of the design of the programme allow for the evaluation?

Figure 6: Evaluability Assessment: Plausibility

Plausibility	Yes	No
Intervention activities should reasonably lead to intended outcomes (Source: Theory of Change)		
It is important to articulate the various activities and components, how they interact and how they are expected to achieve the expected change.		
Depending on the extent of gaps and weaknesses in the intervention logic, a decision may be taken that the evaluation may not be warranted		
It may be better to focus on rethinking and redirecting the intervention logic to increase the potential for outcomes		

Figure 7: Evaluability Assessment: Quality of Design

Plausibility	Yes	No
Is the justification for the project realistic?		
Are the objectives of the programmes clear, realistic and commonly understood by the stakeholders		
Are there SMART Performance Indicators?		
Are the performance indicators monitored adequately?		
Is the log frame flexible and responsive to external factors?		
Were there modifications to the intervention logic and why?		

1. **Feasibility**: Feasibility investigates and determines whether we are "measuring what is worth measuring". Feasibility examines whether the project strategy approach, business models, processes adapted and implementation modalities of the project are feasible and appropriate to assess the outcomes and impacts. Is it feasible to measure the impacts?

- Feasibility of measuring what is worth measuring. Not all data are easy to collect. Data that can be easily collected may not be particularly relevant or appropriate to understand the causality.

- The availability of good quality data and an information system for making data available is necessary. Data and an information system which routinely collects information for performance monitoring and the data gaps, are important.

- The timing of evaluation is crucial in determining what is worth measuring. Premature measurement and belated measurement are both inappropriate for decision-making.

- Design options such as experimental design, quasi-experimental design and the sampling approach are important to investigate causal attribution.

Would the evaluation be feasible, to conduct?

Figure 8: Evaluability Assessment: Feasibility

Feasibility	Yes	No
Does the timing of evaluation fit into the programme cycle? (Usefulness of the evaluation at that point of time)		
Present Stage of Project Execution		
The expenditure rate of the project		
The critical mass of activities has been implemented		
Activities remain to be done		
Have the building blocks of the programme if any, been previously evaluated?		
Can external factors (political/climatic/ security) hamper the evaluation?		
Is the budget sufficient for the evaluation exercise envisaged?		
Are the key stakeholders available for interviews in the field and the headquarters during the planned evaluation time period?		

3. **Credible:** Are the results of the programme verifiable, based on the planned data collection and information system?
   - 3.1 Will baseline data be available to track changes?
   - 3.2 Is the monitoring data planned to be collected on a regular basis against performance indicators?
   - 3.3 Does the present stage of the execution of the programme allow for evaluation?

Figure 9: Evaluability Assessment: Credibility

Availability of Data – Results Verifiable	Yes	No
Baseline data		
Programme documents		
Progress Report		
Review (Tripartite Review report)		
Independent Project evaluation		
Minutes of the meeting		
Studies		
Presentations		
Mission Report		
Agreements and MOUs		
Financial documents—Budget revision		
Any other document relevant to evaluation		

4. **Utility:** The aspect of utility examines whether the project considers the information needs of the programme management, stakeholders and policy-makers and whether it would be useful and used by the relevant authorities for programme improvement and decision-making. Would an outcome or impact evaluation be useful and used? Utility focuses on the following aspects:

    - Considering the stakeholders' needs and expectations: An evaluation should be designed to meet the requirements of the information needs. Different stakeholders have different information needs. A single evaluation may have to satisfy multiple needs.

    - It is necessary to understand who needs what information? When? And for what use or users? Prioritising information needs and expectations is important. Based on the purpose, Key Evaluation Questions (KEQs) should be designed and should include underlying assumptions.

    - Key performance indicators that are relevant and aligned with the "Theory of Change" should be designed for measuring achievements. Quantitative, qualitative and mixed indicators disaggregated by sex, ~~gender~~, age and ethnicity with baselines, should be made available. In the examination of feasibility, it is important to measure outcomes and impacts or intended changes, given the ground realities and resources available.

    - It is a waste of time to conduct an evaluation of which the findings are not likely to be used. Insufficient demand limits the scope for using the findings for decision-making. Lack of credible evidence due to inadequate resources and inappropriate timing may also limit the use of evaluation findings.

    - Weak, non-actionable recommendations may lead to non-utilisation of evaluation findings. In such a situation, it may be appropriate "not to proceed" with the evaluation.

Therefore, the EA will look into very broadly four key questions:

i. Is it plausible to expect the impacts? Does the quality of the design of the programme allow the evaluation?
ii. Is it feasible to assess or measure the impacts?
iii. Are the results of the programme credible and verifiable, based on the data and information collection system?
iv. Will the findings of the evaluation be useful and used? Would the evaluation be feasible, credible and useful?

The EA will also recommend improvements to the evaluation design, Key Evaluation Questions (KEQs), including specific evaluation questions, methodology, tools and techniques where appropriate. The EA should be conducted before the mid-term evaluation so that the evaluation can better document the relevance of the programme, the coherence of the intervention, the efficiency of its implementation, effectiveness, impact and sustainability. The Evaluability Design Matrix of the ILO Supported Local Empowerment through Economic Development (LEED) is shown below:

Figure 10: Evaluability Design Matrix—LEED Project

Criteria / KEQs	Indicators	Data Collection Methods	Source of Information
Plausibility Is the quality of the project design and "Theory of Change" relevant and appropriate to achieve the intended outcomes and project objectives, given the underlying assumptions and contextual factors?	Degree of stakeholder perception of the quality of design and TOC is well articulated to achieve outcomes	Desk Review on TOC and discussions and key informant interviews; discussion with policy-makers, project staff and provincial govt. officials including sector experts, representatives of beneficiaries including co-ops and private sector representatives	Degree of stakeholder perception of the quality of design and TOC is well articulated to achieve outcomes
Feasibility Whether the project strategy, approach and business models are implementable and whether a sound results-based M&E system with data collection and MIS is designed and implemented, to measure what is worth measuring Is the timing of the evaluation appropriate in terms of political, security and other considerations, including stakeholder availability, to undertake meaningful evaluation?	High potential sectors are selected and are appropriate to the context and well-designed. Inclusive and sustainable business models in place Results based M&E system with MIS and data collection mechanism in place with gender and disability disaggregated data. Appropriateness of timing in terms of political, security and availability of stakeholders and other considerations and to ensure effective policy feedback	Key informant interviews with policy-makers, Provincial officials, co-operatives, experts, Chambers, representatives of the private sector and beneficiaries Desk Review including Review of Results Framework with KPIs and Baselines Key informant interviews including policy-makers, programme officers, senior government officials at National, Provincial District and Divisional offices. Discussion with officials of the NPD and the National Budget Department, PTF and private sector representatives, Development partners, ILO and other concerned agencies and beneficiary representatives, private sector representatives and MEMEs Interviews with policy feedback team members of the provincial project steering committee	Responses and perceptions of primary and secondary stakeholders and beneficiaries, including the users of the M&E Annual Progress Reports, M&E Appraisal, Results Framework and Mission Reports Responses and perceptions of primary and secondary

Criteria / KEQs	Indicators	Data Collection Methods	Source of Information
Credibility Is there adequate, verifiable, evidence- based data and information including results reporting and monitoring data and a tracking system available to undertake credible evaluation?	Availability of data, information and verifiable evidence and authentic information Willingness of primary and secondary stakeholders, including beneficiaries to provide information and express frank and free opinions.	Desk Reviews, discussions and interviews with project staff, ILO, Provincial Councils and Primary and Secondary Stakeholders including co-ops, MSMEs and private sector processing centres	Review of progress reports, mission reports, minutes, KPIs, studies, evaluation reports, MOUs including data and information at grass-roots level with co-ops, MSMEs, farmers and private sector processing centres.
Utility Ability of the evaluation to meet the multiple information requirements of different stakeholders Is there demand for evaluation Information and Are the findings and actionable Recommendations are likely to be used effectively to improve programme intervention and policy feedback	Perceptions of primary and secondary stakeholders about the usefulness and use of evaluation. Perceptions of primary and secondary stakeholders, Management Response and Improvement Plan in place.	Key Informant Interviews with primary and secondary stakeholders and members of the provincial project steering committee and users of the EA. Perception survey of policy-makers and users of the evaluation findings	Responses of policy-makers, provincial project steering committee members, programme managers, evaluation commissioners and implementing agencies (including the users of the EA) Management Response and Improvement Plan

5.4 Recommendations:

Based on the evaluability assessment, the following recommendations will be made.

1. Proceed/Not to proceed with the Evaluation

- Evaluation should take place: The programme is ready for evaluation. The programme is deemed evaluable and is ready for evaluation. It is also likely that the Evaluability Assessment will suggest some parameters both criteria and performance measures for evaluation.

- Evaluation should not take place: The programme is not ready for evaluation. It is not ready because of fundamental planning and implementation issues. If the programme is not ready for evaluation, the EA should recommend action to be taken to bring the interventions to an appropriate level.

2. Increase Readiness: If the programme has potential to improve readiness, then the Evaluability Assessment will recommend proceeding, but after addressing the critical areas. These may include:

- Strengthening the project design and Theory of Change and revise the project activity.
- Modify project activities to have a greater chance of reaching the stated project objectives (programme modifications).
- Improving the project or activity management.
- Improve measurement by revisiting the data and information system.
- Strengthen the commitment to evaluation.

Figure 11: Summary of Decision Support arrangements

Decision Support	
1. Proceed with the impact evaluation	No major barriers exist
	Clear purpose; demand for evaluation; key Evaluations questions have been agreed; M&E framework in place and strategy in place to share and use the findings
2. Proceed but address the critical issues	Evaluation is feasible in the near future
	Different stakeholder expectations; the number of evaluation questions exceed the budget and timeline; prioritising the evaluation question is possible and evaluation is feasible; the M&E framework is available but gaps need to be addressed; strategy in place to use the findings but needs improvement
3. Do not proceed with evaluation	Critical barriers cannot be addressed easily, in time
	Demand for impact evaluation is lacking; goals and objectives are not clearly defined; designs/Theory of Change needs major improvements or does not exist; the M&E framework is non-existant or needs major improvements; prioritising the evaluation questions is not easy; strategy not in place for use of findings

5.5 Evaluability: Improve the Quality of the Evaluation

1. Evaluability assessment occurs before an evaluation, with the intent to recommend whether or when the evaluation should take place and how to maximise its value. It improves evaluation quality and use. It helps to use robust evaluation design to reduce the risk of irrelevant and invalid findings.

2. A small cost from the evaluation budget will prevent the waste of valuable time and resources on premature or inappropriately designed evaluation. The value for money of an evaluation is determined based on the use of the evaluation findings and its ability to influence decision-making.

3. Evaluability Assessments help to decide (a) whether to proceed with the evaluation; (b) proceed, but after addressing the critical issues in the near future or (c) not to proceed with the evaluation

4. Evaluability assessment helps
   - To improve intervention design.
   - To improve the design of the M&E system.
   - Improve the use of evaluation findings.
   - Recommends how to re-design the evaluation to ensure that the stakeholders' needs and expectations are captured.
   - Decide whether the planned evaluation should take place now or later.

If there are gaps in the intervention logic, it may not warrant an immediate evaluation. Instead it may be better to focus on rethinking and redirecting the intervention logic to increase its potential for effectiveness.

CHAPTER 6

Step 3: Terms of Reference and Selection of Evaluators

6.1 Purpose, Features and Contents of the TOR for Evaluation

The TOR provides the background for evaluation, defines its purpose and objectives, the scope of work, the composition of evaluation team and the timing of evaluation. It is important to prepare the Terms of Reference for the evaluation and this Chapter provides an outline on key matters to be covered in the TOR. The Commissioner of Evaluation (COE) prepares the TOR, once the decision is made to proceed with an evaluation. The TOR articulates the requirements and expectations of the evaluation and is used to manage the evaluation. The TOR for evaluation will include the following salient aspects and contents:

6.1.1 Information on the Title of the Evaluation

- Name of the Project to be evaluated.
- Type of Evaluation (on-going ex-post or impact evaluation).
- Commissioner of the Evaluation (Name/designation/organisation).
- Date of Evaluation (Proposed date of commencement and proposed date of completion).

6.1.2 Introduction/Rational for the Evaluation of the Intervention

- Development of the intervention, its scope, beneficiaries, duration and cost;
- Need for the intervention;
- Its activities, outputs, outcomes and impacts;
- The main Theory of Change (TOC) that underpinned the intervention;
- The partners and stakeholders;
- The reasons/rationale for the evaluation of the intervention at this point in time and any decision that may be made using the results of the evaluation.

6.1.3 Focus of Evaluation

a) The Purpose of Evaluation

This section should answer the question: What is it that we want to understand about the intervention?

Example:

- Is the logic of the intervention/design robust and likely to work?
- How did the beneficiaries' life change as a result of this intervention?
- Is it conducted for the accountability for results or learning/knowledge management/decision making—it may include scaling-up, scaling-down, termination or re-designing of the intervention.

b) Key Evaluation Questions (KEQs)

- The Key Evaluation Questions (KEQs) are high level questions that an evaluation is designed to answer and should reflect the stakeholders' needs and interests. It provides the direction and foundation of the evaluation.
- This section indicates detailed evaluation questions which are being asked within the overall core questions and for which answers are sought.

- It will be based on OECD/DAC evaluation criteria namely: relevance, coherence, efficiency, effectiveness, impact and sustainability—R/C/E/E/I/S.
- The Service Provider is expected to conduct the scoping session with key stakeholders, finalise the evaluation questions and submit the analytical framework for the evaluation and Evaluation Design Matrix

C) Intended Users and Stakeholders of the Evaluation

This should indicate key potential users of the evaluation findings and how they will use it.

Figure 12: Potential Users of the Evaluation and Use

Potential Users of the Evaluation	How they will use it
Commissioner of Evaluation	Initiate Management Responses and develop an Improvement Plan
Primary intended users—Those who make decisions on the basis of the evaluation findings.	
Policy-Makers and Programme Management staff	Scaling-up, scaling down, termination, modification and amendments to the programme
Department of National Planning	Learning Lessons to improve future projects
Department of National Budget	Evaluation to be used in the allocation of budgetary resources
Parliament, including the Auditor General/Committee on Public Accounts (COPA), Committee on Public Enterprises (COPE)	Performance Accountability
Other stakeholders	Those who will be affected by interventions made during and after the evaluation

d) Scope of Evaluation

- This indicates the components, sector, thematic areas and geographical areas to be covered and the matters outside the scope of this particular evaluation.
- In addition, Evaluation should include the adoption of a human rights-based approach, gender equality, non-discrimination and environmental sustainability.

6.1.4 Evaluation Design and Methodology

This section covers the approach, design and methodology used by evaluation team. The methodology may include quantitative, qualitative and mixed methods.

- Desk Research: Document Review, Review of Project Appraisal Document (PAD) Project Completion Report (PCR) and Project records
- Meetings with Stakeholders
- Key Informant Interviews: National Project Director, Steering Committee Members, Donors and Focal Persons
- Conduct Field Missions—Interviews, Focus Group Discussions, Community Interviews and surveys.
- Conduct Stakeholder Consultation Workshops
- Econometrics and Statistical Analysis
- Case Studies and Counterfactual Analysis
- Sample Surveys, including sample size

6.1.5 Main Deliverables

- An Inception Report describes the conceptual framework to be used for undertaking the evaluation and it includes the following milestones and deliverables.
- Development of a draft Theory of Change
- Proposed methodology
- Evaluation questions, data required, data collection methods and the Evaluation Design Matrix
- Detailed Work Plan of the evaluation, including deliverables and milestones
- Key stakeholders to be interviewed and tools to be used
- Set out the outline for final report. This should be in the following prescribed format:
  - Cover page with key project dates, title, donor, start and end date, budget, management arrangements, geographical coverage including evaluation mission, name of evaluators and date of commencement and completion of the evaluation report
  - Table of contents
  - Acronyms
  - Executive Summary
  - Background of the Project and Intervention Logic
  - Purpose, Scope and client of evaluation
  - Methodology and Evaluation Questions
  - Review of Planning and Design, Implementation and Post-implementation aspects.
  - Overall Assessment
  - Key Issues, Lessons Learnt, Follow-up actions and Recommendations
  - Annexures—List of Interviews and Meetings held.
- Preliminary Findings to be presented at the stakeholder's workshop
- Draft Evaluation Report: Follow the 1/5/25 structure and prescribed format
  (A 1-page high level executive summary for Ministers/Secretaries of Ministries; A 5-page executive summary for senior officers and a 25-page report for operational managers)
- Workshop with stakeholders to discuss the draft report and emerging findings
- The final Evaluation Report in 1/5/25 structure and in the prescribed format
- A Power Point Presentation of the results with photographs. Report to be placed in the website.
- The evaluation team may also be asked to assist the Commissioner beyond the evaluation report, especially in the development of the Management Response and Improvement Plan

6.1.6 Payment Milestone—The Payment Schedule is set out below and changed where appropriate.

Figure 13: Payment Milestone/Schedule

Deliverable	% of Payment
Approval of Inception Report	10%
Submission of Draft Theory of Change with LFA	10%
Submission of Literature Review	10%
Submission of Fieldwork Report supported by Field Data Collection & Instruments	10%
Submission of Draft Evaluation Report 1/5/25	30%
Stakeholder Workshop on Draft Report/Final Report	20%
Submission of Power Point Presentation	10%

6.1.7 Budget—The National Evaluation Plan provides an indicative budget for this assignment. The service provider should make sure the scope is realistic for the amount available.

6.1.8 Management Arrangements

- An Evaluation Manager from the Commissioning Department will liaise with the service provider on a day-to-day basis on routine matters
- A Steering Committee with the Commissioning Department and other Implementing Agencies will guide the study.
- A Reference Group—will comprise of an inter-ministerial committee of Evaluation Heads of the relevant Ministries/Development Partners, eminent persons and some subject area specialists who will provide strategic guidance in the planning and conduct of the evaluation.
- A Peer Review Arrangement to ensure quality.

6.1.9 Evaluation Team

- The Evaluation Manager is responsible for the overall management of the evaluation, quality control and liaison with the client.
- The Evaluation Specialist brings specialised knowledge on evaluation methodology.
- The Sector Specialist brings in-depth knowledge of the sector in undertaking the evaluation.

6.1.10 Evaluation Competency Framework

Competency is defined as the "Cluster of related knowledge, skills, ability and other requirements necessary for successful job performance". To ensure high quality evaluations, five areas of knowledge, skills and abilities are required.

1. Professional Foundation in evaluation
2. Technical Evaluation Skills
3. Evaluation Management Skills
4. Inter-personal skills
5. Promoting a culture of learning from evaluation

6.2 Selection of Evaluators

The evaluator or service provider should be invited through a call for expression of interest from qualified evaluation consultants to conduct the independent evaluation. The procurement procedure of the Government of Sri Lanka should be followed in selecting the evaluation consultant.

The call for Expression of Interest should generally request candidates to submit the following information

1. The candidate's skills, qualification and experience
2. Previous evaluation assignments undertaken and other involvements in evaluations as a member of the steering committee, management committee, reference group or peer review
3. Candidate's curriculum vitae
4. A statement confirming that the candidate has no previous involvement in the planning, implementation and post-implementation of the relevant project identified for evaluation.
5. The professional fee and other related information

While selecting the evaluator or consultant, due consideration should be given to skills, professional competences and evaluation experience. These are generally identified in the UNEG as the Evaluation Competency Framework and includes five clusters of evaluation-related skills, knowledge, ability and other

professional exposure. These skills, competencies and knowledge are covered in 5 areas namely: i) Professional Foundation; ii) Technical evaluation skills; iii) Management skills; iv) Interpersonal skills and v) Promoting a culture of learning for evaluation.

Figure 14: Evaluation Competency Framework (Source: UNEG)

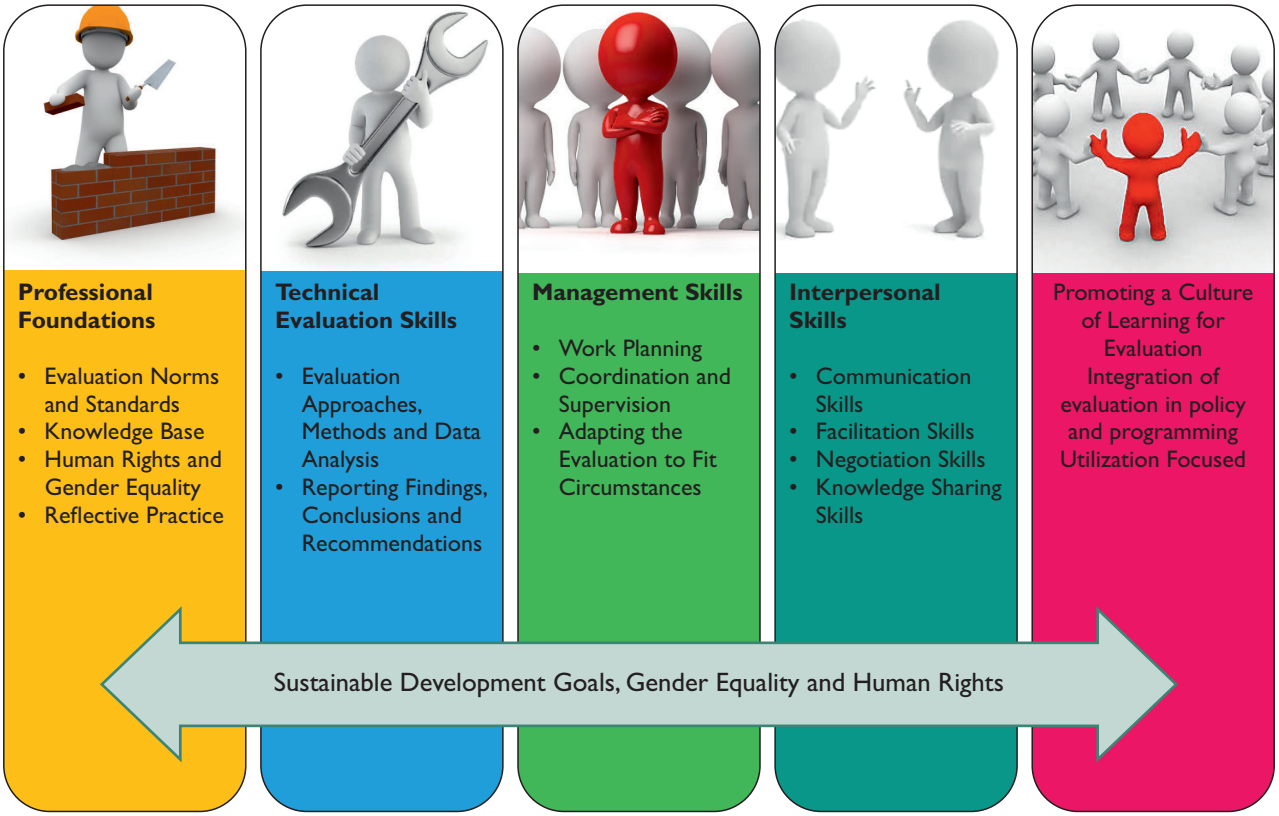


Figure 15: Evaluation Competency Framework: Cluster of Knowledge, Skills and Ability required

Competency Area	Demonstrated Ability
Professional Foundation	• Ethics and integrity (Confidentiality, independence and impartiality) • Evaluation norms and standards • Knowledge base: Understands the evaluation theory and practice and the current issue in evaluation • Integration of human rights and gender perspectives in the evaluation • Reflective practices: Support the professional development of other evaluators and link with the evaluation network.
Technical Evaluation Skills	• Norms and standards • Quality Assurance Mechanism, Quality standards • Evaluation purpose and design: Identification of purpose, who will use the findings, types of decisions based on findings, evaluation design, best practices, and technically and culturally appropriate design • Knowledge of Data Collection and Analysis Methods Knowledge of evaluation design for complex projects, Evaluation approaches, methods and data analysis, quantitative and qualitative methods and analysis, data collection and analytical skills, interpreting data, ability to use multiple and mixed methods, biases and data reliability, credible findings based on data analysis, use of diverse sources.

Competency Area	Demonstrated Ability
Professional Foundation	• Reporting findings and recommendations - Credible conclusions and findings based on evidence - Clear, realistic, feasible and actionable recommendations - Ability to disseminate and use evaluation findings - Awareness and knowledge of integrating findings into future programs
Evaluation Management Skills	• Work Planning: Determining the scope of evaluation, the Evaluation Plan and budget and managing resources • Coordination and supervision - Supervising the work of the evaluation team, planning and processes - Efficient and effective work delivery and meeting deadlines - Engagement of all stakeholders and ensuring quality - Assess the risk and contingency plan • Adapting the evaluation to fit circumstances - Identify conflict issues and problem-solving strategies - Conflict Resolution skills and managing conflicts • Respect local customs, religious beliefs, gender roles, ethnicities and be sensitive to disabilities.
Inter-Personal Skills	• Communication Skills - Written, oral and listening skills - Communicate complex technical evaluation in an understandable form to non-experts - Adapt communication to different audiences - Ability to actively listen to others, excellence in asking questions, write user-friendly evaluation reports and presentations • Facilitation Skills - Ability to employ mechanisms to engage diverse beneficiaries and users - Able to integrate diverse perspectives - Excellent in meeting and facilitation skills to engage stakeholders • Negotiation skills - Negotiate evaluation priorities—excellent negotiation skills to ensure high quality evaluation • Knowledge Sharing Skills • Ability to develop a communication plan for dissemination of findings. • Build network and partnerships with various stakeholders • Effective engagement with Executive Board and stakeholders - Sharing knowledge
Promoting a culture of learning for evaluation	• Engage users and beneficiaries in evaluation • Integration of evaluation into policy and programme • Utilisation Focused Evaluations and promote learning • Knowledge of mechanisms to promote learning • Promote evaluative thinking • Knowledge of utilisation-focused evaluations and approaches

6.3 Personality of Evaluator

An accountability-oriented evaluator is an independent outsider who plays the role of a watchdog. Lessons Learning evaluation is a participatory process where insiders are involved with a facilitator. The personalities of evaluators have evolved over time and are identified as four generations:

- Technician (First Generation)—First-generation evaluators are recognised as measurement-oriented personalities using science-dominated, rigorous methodologies. Evaluators are considered technical experts and measurement oriented, undertaking tests to measure effectiveness.
- Describer (Second generation)—The second-generation evaluators were recognised as "Describers" analysing the strengths and weaknesses of a project.
- Judge (Third generation)—The third-generation evaluators are recognised as judgment-oriented personalities and follow standards as objective evaluators.
- Negotiator (Fourth Generation)—The fourth-generation evaluators were recognised as "Negotiators" of diverse options and issues, with the participation of all stakeholders. The Evaluator is considered a collaborator, learner, reality-shaper and change agent.

A good Terms of Reference and the selection of a quality evaluator will contribute significantly to the success of an evaluation.

CHAPTER 7

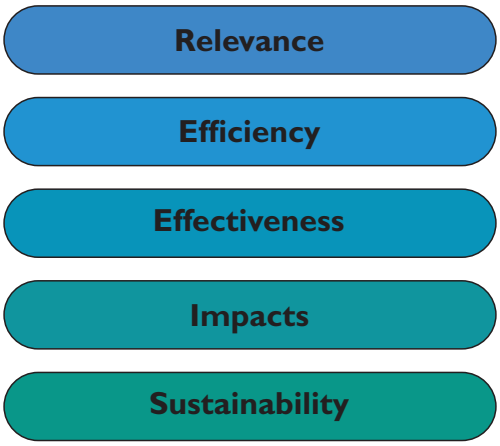
Step 4 and 5: Evaluation Criteria, Key Evaluation Questions and Scoping Session

7.1 Evolution of Evaluation Criteria

Prior to DAC Criteria, Development interventions were assessed on the basis of outputs; use of funds and Economic Rate of return (ERR).The Evaluation Criteria were first set out by the OECD/DAC in 1991.These are the most widely used and they focus beyond inputs, activities and outputs towards results. Five OECD/DAC Evaluation criteria have been recommended by the OECD and adapted in Sri Lanka as standard yardsticks for the evaluation of development interventions.

The purpose of the evaluation criteria is to enable the determination of merit, worth or significance of the intervention. Each criterion is a different lens or perspective through which the intervention can be viewed. Together, they provide a comprehensive picture of the intervention. Together, they describe the desired attributes of interventions: All interventions should be relevant to the context, coherent with other interventions, achieve their objectives and desired results in an efficient way and have positive impacts that last. They can be used to look at processes to observe how change happens as well as the results of what changed.

Figure 16: OECD/DAC Evaluation Criteria



DAC Evaluation Criteria helped evaluation to reach beyond input-output thinking (Robert Picciotto).The DAC Criteria is the right first step to design the right evaluation questions and as such, they are considered vital and important in planning, design and undertaking evaluations.These criteria are widely used by partner countries, development partners and INGOs.

Relevance, Efficiency, Effectiveness, Impact and Sustainability (REEIS) is identified as the Big Five criteria used as "gold standards" in evaluation.They are not imposed but picked-up by many different actors.

The DAC network on development evaluation adapted the DAC evaluation criteria in the light of changes in the development landscape and the adoption of the 2030 agenda and the Paris Climate Agreement.

In 2019, the criteria were improved by the OECD/DAC, in consultation with all stakeholders, with improved definitions, capturing new criteria "coherence" to better capture synergies, linkages, partnership dynamics and complexity and responded to cross-cutting focus areas such as equity, gender equality and "leave no one behind" imperatives.

7.2 OECD/DAC Evaluation Criteria

Since 2019, the OECD/DAC has identified 6 criteria, namely: relevance, coherence, efficiency, effectiveness, impact and sustainability.

(a) Relevance: Is the intervention doing the right thing?

The "relevance" examines the extent to which the development intervention meets the needs and priorities of target groups and is aligned to the national development plan of the country and the Sustainable Development Goals (SDGs). Relevance examines how far the intervention fulfils the national priorities set out in the national development policies and meets the priorities, aspirations and expectations of the beneficiaries. It also examines the question of "Whose reality counts?". It is the priority of the beneficiaries, rather than the donors.

Relevance also implies the question of whether the activities, components and outputs are likely to achieve the expected purpose and goals of the project. From an intervention context, it examines whether the "Theory of Change" and "intervention logic" are robust. Similarly, the balance between various activities and components is of importance to achieve results. The KEQs in relation to relevance criteria may be defined as shown below:

Key Evaluation Questions

- Is the logic of the intervention and design robust and likely to work?
- Have the right things been done?
- Are the activities, outputs and components relevant, appropriate, balanced and consistent to achieve the objectives?

The aspect of "Relevance" focuses primarily on the goals and priorities set by country governments, SDGs, targeted population (intervention impactees) and development partners.

(b) Efficiency: How well are the resources used?

The efficiency examines the extent to which the cost of the development intervention is justified by its results, taking into consideration the alternatives. In other words, it examines whether the intervention delivered or is likely to deliver the results in an economic and timely way. Efficiency is more concerned about cost-effectiveness than least cost. It examines the delivery of outputs in time, within cost and with quality. Time overrun and cost overrun should be avoided and appropriate efficient implementation modalities and institutional arrangements are vital. Efficiency also examines how well the intervention was managed. Efficiency is a measurement of the relationship between "inputs" and "outputs". The KEQs in relation to efficiency criteria may be defined as follows:

Key Evaluation Questions

- Is the intervention being executed and implemented as specified and if not, why?
- Is the project implemented on time, within cost and in compliance with the quality requirements?
- Was the project implemented in the most efficient way, compared to alternatives?

Consideration should be given to cost-effectiveness, value for money and non-**monetary cost**s such as social and environmental costs **and be** factored and replaced to 'least cost'. In focusing on **cost-effectiveness, it is important to examine creative alternative activities, interventions and strategies.**

(c) Effectiveness: Is the intervention achieving its purpose and objectives?

Effectiveness examines the extent to which an intervention has achieved or is expected to achieve its intended outcomes and impacts. Effectiveness measures the relationship between "outputs" and "outcomes".

- Effectiveness focuses mainly on the needs of the beneficiaries and examines the outcomes or purpose.
- It is also necessary to ensure that effectiveness is not included under impacts and that there is no overlapping.
- To what extent were the objectives achieved or likely to be achieved?

The KEQs in relation to effectiveness may be defined as follows:

Key Evaluation Questions:

- Are the programme benefits reaching their targeted beneficiaries? And are the outcomes realised?
- What results can be attributed to the intervention?

(d) Impact: What difference is the intervention making?

The term "impact" examines the direct and indirect, intended or unintended, positive or negative long-term results of an intervention. It is the relationship between "purpose" and "goals". It examines the economic, social, environmental, institutional and other cross-cutting aspects such as gender, human rights etc. It examines the significance of the potentially transformative effects of the intervention, including the social, environmental and economic impacts.

The KEQs in relation to "Impact" may be defined as follows:

Key Evaluation Questions:\

- How have the beneficiaries' lives changed as a result of the intervention?
- What transformational changes are made by the intervention?

Unintended impacts need special attention. Both positive and negative impacts need to be considered.

(e) Sustainability: Will the benefits last?

Sustainability examines the continuation and longevity of benefits from a development intervention, even after the termination or cessation of development assistance. It is the ability of the intervention to continuously deliver services uninterruptedly, even after the withdrawal of technical, managerial and financial support from the programme intervention. It includes the examination of financial, economic, social, environmental and institutional capacities of the systems needed to sustain net benefits over time. It analyses resilience, risks and potential trade-offs.

Sustainability also examines financial and environmental sustainability. Financial sustainability involves arrangements in place for operation and maintenance, self-income generation and cost recovery. It is important to ensure smooth handing-over and taking-over with institutional arrangements to sustain the service delivery. Sustainability is also extended to **retrospective** sustainability (how sustainable it has been). Other

contributory elements for sustainability include policy support, political support, cultural appropriateness, adequacy of technology, institutional capacity, environmental and social aspects and community participation.

Key Evaluation Questions

- The extent to which the net benefits of the intervention continue or are likely to continue after the termination of the development intervention.
- What type of strategies and systems are in place for the operation and maintenance of the facility, including financing and institutional capacity.

(f) Coherence: How well does the intervention fit?

The compatibility of the interventions with other interventions in the country, sector or institutions is important for coherence. The perspective of development is wider and more complex in an SDG era and it needs to look at inter–connectedness, inter-linkages, synergies and trade-offs. Achieving one SDG goal may have impacts on other goals. Coherence looks at the synergies and inter-linkages as well as the consistency of the intervention with other interventions. This includes complementarity, harmonisation and coordination with others, while avoiding the duplication of efforts.

Example: Achieving SDG Goal No. 8 of economic growth through the extensive use of fertiliser and pesticides for agriculture production may adversely effects the Goals of 13, 14 and 15 of sustainable environment and climate change, including health and well-being. Therefore, the application of fertiliser and pesticide as a target to improve agriculture projects needs to be evaluated after taking into consideration the environmental sustainability and health and wellbeing of the community.

Figure 17: New OEC/DAC Evaluation Criteria

Relevance	Is the intervention doing the right thing?
Coherence	How well does the intervention fit?
Efficiency	How well are the resources used
Effectiveness	Is the intervention achieving its objectives?
Impact	What difference is the intervention making?
Sustainability	Will the benefits last?

Cross-cutting issues such as poverty, gender, environment and human rights are taken into consideration as part of the evaluation criteria.

7.3 Key Evaluation Questions and the Theory of Change

Developing the Key Evaluation Questions (KEQs) is the crucial first step in designing an evaluation. These identify what the stakeholders want to know about the programme through the evaluation. Developing a good set of evaluation questions is very important because the questions will drive what type of information (data) you collect.

The most important step in development evaluation is to formulate key evaluation questions (KEQs). The key evaluation questions are usually developed based on two important points:

(i) Theory of Change (TOC)
(ii) Evaluation Criteria

7.3.1 Theory of Change

"Programme theory" serves as the foundation for evaluation questions. The "theory of change" (TOC) articulates the linkage between intended inputs, planned activities and expected results. It sets out and articulates the various activities and components (such as mid-day meals delivered to needy schoolchildren), how they interact and how they are expected to achieve the desired change (Improved dynamism, concentration and learning outcomes of schoolchildren). It shows the logical linkage between activities (eg. Provision of fishing boats and nets to vulnerable groups) and intended impacts (eg. Improved income and livelihoods of a vulnerable fishing community) and sets out the context and the underlying assumptions that influence the results chain.

Theory of change is defined as a "set of beliefs" about how and why an initiative will make a change. It shows the relationship between intended inputs, planned activities and expected results. It is a mental model that visualises and maps out what activities are most important to bring about the change.

The Theory of Change is a mental model that visualises and maps out which activities are important to bring about the desired change. It shows the causal relationship between inputs and activities with outcomes and impacts. The Theory of Change describes the causal mechanism of how activities and outputs will result in expected outcomes and impacts. The Theory of Change is defined as a "set of beliefs" about how and why an initiative will make a change. It shows the relationship between intended inputs, planned activities (Rehabilitation of irrigation schemes and water management practices) and expected results (increased agricultural production and farmer's income).

Figure 18: The Theory of Change

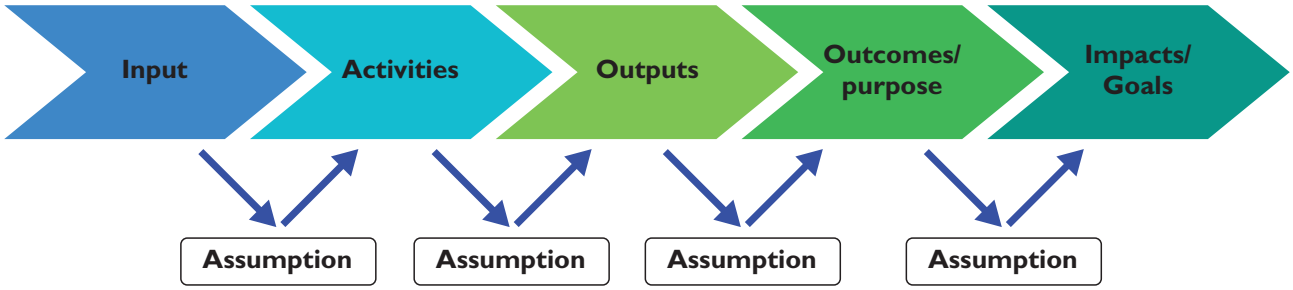
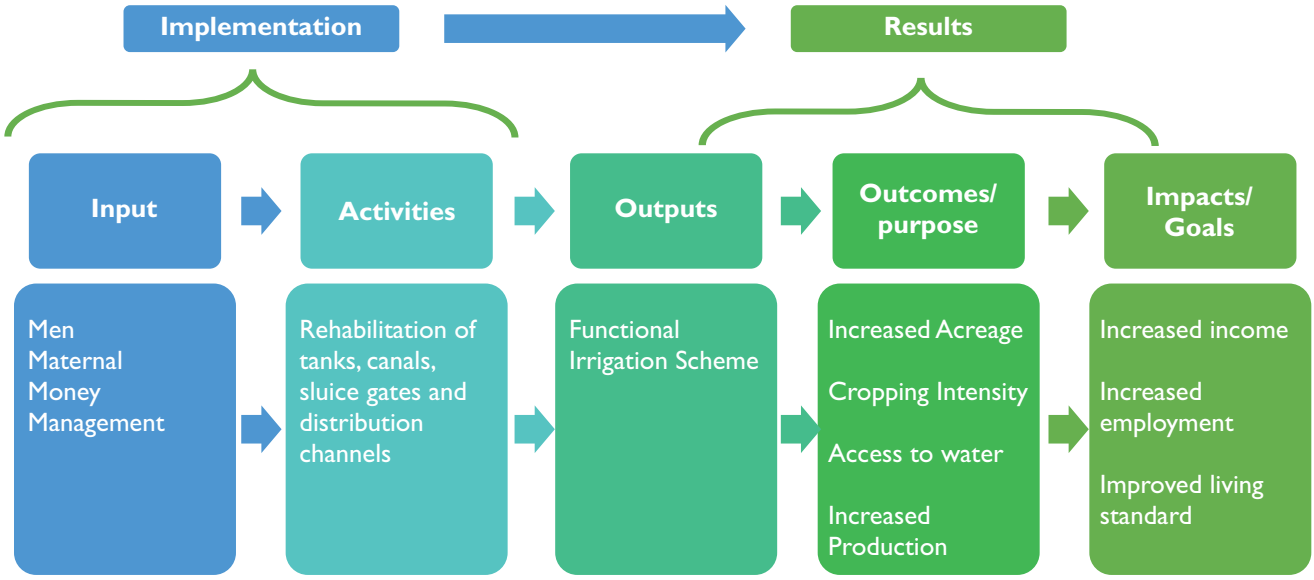
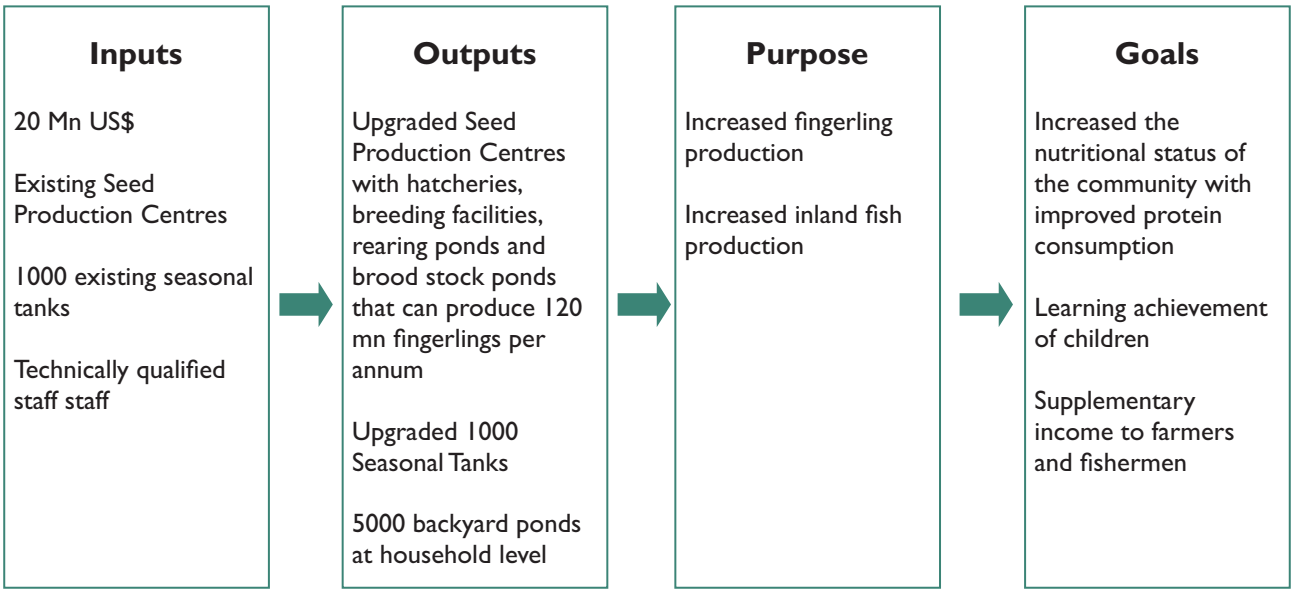


Figure 19: Theory of Change National Irrigation Rehabilitation Project (NIRP)



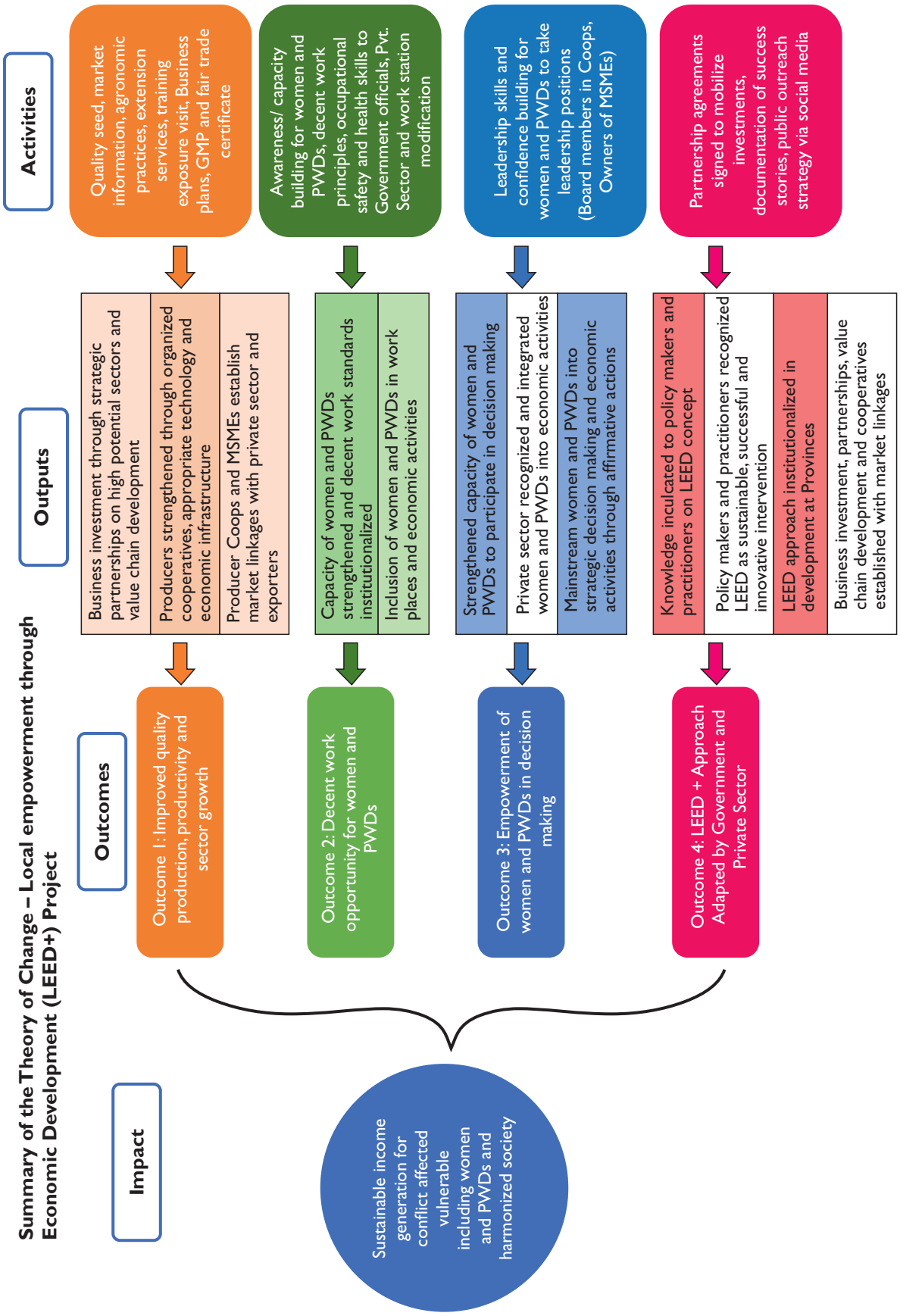
The uncertainty of the project is explained by the assumptions at each level. If the expected assumptions fail to realise, it will adversely affect the results chain. In the National Irrigation Rehabilitation Project, irrigation schemes were rehabilitated and water management practices were improved (activities and outputs) to achieve improved agriculture production and farmer's income (outcome/impact). If the farm families are engaged and interested in white collar jobs, it is impossible to get them to farm and as a result, agriculture production and income may not be improved as expected. Therefore, these "killer assumptions" need to be internalised by incentivising farmers and by promoting and recognising farming.

Figure 20: Theory of change—Inland Fish Production Project



The Theory of Change is best developed through a combination of the Desk Review of the project document, literature review of a similar programme and stakeholder discussions about how they think the project works. It involves an iterative participatory process. The 'Theory of Change' can be represented in the form of a logic model, results chain, log frame, programme theory or outcome mapping. 'Outcome mapping' is a particular approach to develop a Theory of Change which is suitable when a programme does not directly produce the intended results but works through influencing the behaviour of people. Evaluators ask questions about whether the Theory of Change is convincing. If anything important is missing and if the programme has not worked, the TOC can help us to identify whether this is due to (i) failure in implementation or (ii) because the Theory of Change does not work.

If the programme has worked, the TOC can help to identify what is needed to repeat the success at another time or another site. Are the activities and outputs appropriate, necessary and sufficient to achieve the outcomes? Are there are any important activities or outputs missing for achievement of the outcomes? Outputs are not just products; they are the building blocks needed to achieve outcomes. It is necessary to understand whether the root cause of the problem is being addressed by the intervention.


Summary of the Theory of Change – Local empowerment through Economic Development (LEED+) Project
Activities
Outputs
Outcomes
Impact
Quality seed, market information, agronomic practices, extension services, training exposure visit, Business plans, GMP and fair trade certificate
Awareness/ capacity building for women and PWDs, decent work principles, occupational safety and health skills to Government officials, Pvt. Sector and work station modification
Leadership skills and confidence building for women and PWDs to take leadership positions (Board members in Coops, Owners of MSMEs)
Partnership agreements signed to mobilize investments, documentation of success stories, public outreach strategy via social media
Business investment through strategic partnerships on high potential sectors and value chain development
Producers strengthened through organized cooperatives, appropriate technology and economic infrastructure
Producer Coops and MSMEs establish market linkages with private sector and exporters
Capacity of women and PWDs strengthened and decent work standards institutionalized
Inclusion of women and PWDs in work places and economic activities
Strengthened capacity of women and PWDs to participate in decision making
Private sector recognized and integrated women and PWDs into economic activities
Mainstream women and PWDs into strategic decision making and economic activities through affirmative actions
Knowledge inculcated to policy makers and practitioners on LEED concept
Policy makers and practitioners recognized LEED as sustainable, successful and innovative intervention
LEED approach institutionalized in development at Provinces
Business investment, partnerships, value chain development and cooperatives established with market linkages
Outcome 1: Improved quality production, productivity and sector growth
Outcome 2: Decent work opportunity for women and PWDs
Outcome 3: Empowerment of women and PWDs in decision making
Outcome 4: LEED + Approach Adapted by Government and Private Sector
Sustainable income generation for conflict affected vulnerable including women and PWDs and harmonized society

If a Theory of Change does not exist, then the first activity in the evaluation should be to develop a TOC, to understand how a project works. The evaluator should examine the intervention logic and the Theory of Change and then develop the evaluation questions. Evaluation questions should be categorised according to the broad evaluation criteria—relevance, coherence, efficiency, effectiveness, impact and sustainability.

7.3.2 Key Evaluation Questions (KEQs)

Key evaluation questions are high level questions that an evaluation is designed to answer. The KEQs define what the project team seeks to learn and serve to guide the entire evaluation process, inform decisions about what data to collect through which method, how to analyse and interpret it and how to disseminate and report findings. Once the key evaluation questions are agreed it is easier to decide what data is to be collected, how to analyse it and how to report it.

Evaluation questions provide the direction and foundation for evaluation. Without evaluation questions, the evaluation would lack focus and the evaluator would have difficulties in explaining what is to be examined and how and why it is being examined.

Key Evaluation Questions (KEQs) are usually developed and agreed at the beginning of evaluation planning. Asking the right evaluation questions is very important for successful programme evaluation. The evaluation questions should provide the big picture and the critical issues, and help to understand the actionable recommendations. It is necessary to raise whether the **key evaluation question** is very important and worthwhile and will it help to change the way the programmatic intervention works. Trivial questions should not be allowed as they would consume evaluation resources and the results will be a waste of effort. It is essential to have a sound balance between evaluation questions and the time and resources available to answer them.

Figure 22: Evaluation Question in Budget and Time Constraints

The foundation for successful evaluation is based on the importance of key evaluation questions. It is necessary to understand why evaluations are important:

(i) Define the topics the evaluation will investigate.
(ii) Guide the evaluation planning process: Focus of the evaluation, data collection, analysis, reporting and communication.
(iii) Guides the management and the conduct of evaluation, including the reporting on the findings.
(iv) Writing a well thought out, relevant evaluation question is the secret to meaningful evaluation.
(v) KEQs helps the evaluator to see the programme through the eyes of various sectors of the decision-making community, including professionals and citizens.

7.4 Features of Good KEQ

(i) Evaluation questions should be based on the 'Theory of Change' (TOC) and categorised according to the OECD/DAC evaluation criteria.

(ii) KEQs should be small in number, with a maximum of 5-7 main questions. It might also be useful to have some more specific questions under KEQs. The KEQs should address major critical issues and focus on the big picture.

(iii) KEQ should be clear, specific and well defined; focused on the programme or programme components; measurable by the evaluation; aligned with the logic model; linked clearly to the evaluation purpose and reflect the stakeholder's needs and interest.

(iv) Ensure a sound balance between the evaluation questions and the time and resources (budget and human resources) available to answer them.

(v) KEQs should be insightful questions that provide a voice to less powerful stakeholders and would take into consideration the views of all the stakeholders in a balanced manner. The questions should consider the type of evaluation being done, its intended users and uses.

(vi) Active stakeholder engagement is vital in the development of KEQs. Brainstorming evaluation questions is important and should be prioritised according to importance. The KEQ must be relevant to the purpose of evaluation and useful to the decision makers and stakeholders. Based on the findings, the intervention could be expanded, continued, improved or discontinued.

(vii) Develop KEQs that are meaningful, important and feasible to answer, given the available resources and they should provide useful information to the primary intended users and other stakeholders.

(viii) KEQs must be feasible. The data and information should be available to answer the KEQs. The KEQs must also consider the time, resources and complexity of the issues and the availability of data and ability to manage the evaluation.

(ix) KEQ should provide answers based on evidence which can be acted upon (utility). The selection of KEQs should take into consideration the interests of the key audience; contribute to reduce present uncertainty; yield important information; be critical in nature with continuing interest; have an impact on the results chain and be answerable with available data and information.

There are two phases in the formulation of KEQs, namely: the divergent phase and the convergent phase. In the divergent phase, comprehensive lists of questions are developed. In the convergent phase the list of questions is narrowed down, based on the constraints of criticality, priority, programme theory, availability of data, time, budget and resource.

7.5 Step 5: Scoping Session

The Scoping Session on evaluation is usually undertaken to define the breadth and depth of the evaluation study and it covers the purpose, scope and coverage of the evaluation and the limitations. The scoping session is usually conducted with the primary and secondary stakeholders, which includes the Policy-Makers, Senior Government Officials, Programme Managers, the Commissioners of evaluations including the Management Group, Reference Group and Peer Reviewers and the users of the evaluation.

The session should not exceed 50 or 60 participants. It is a session that discusses, brainstorms and conducts dialogue on the road map of the evaluation, with special focus on planning the evaluation. The session will discuss the concept of the project, the Theory of Change and project strategies, revisit the terms of reference and discuss ways to shape the evaluation ideas. The core aspects of the scoping session are to understand the Theory of Change, the logic models, the underlying assumptions and contextual factors and to prioritise the evaluation questions under the evaluation criteria, define key performance measures and indicators, data collection

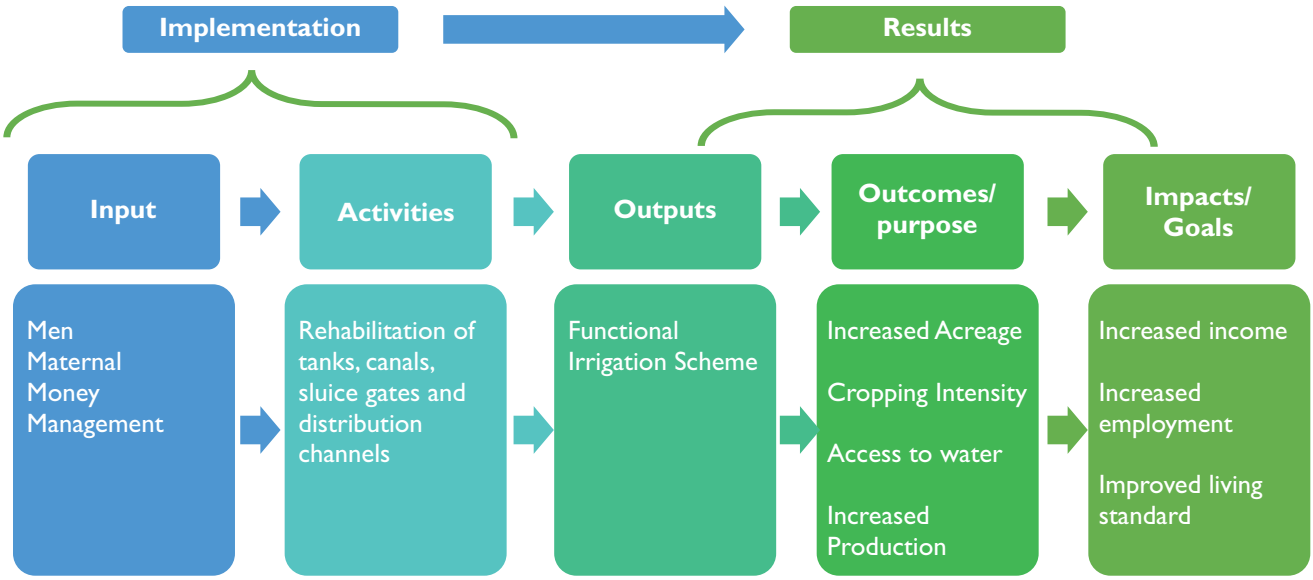
methods, data sources, and research design methods. The scoping session obtains the consensus of the stakeholders on the conceptual framework, evaluation design matrix, logic models and M&E Framework, to meet the expectations of the clientele and primary stakeholders who will use the evaluation findings in the decision-making process.

The objective of the scoping session is to ensure that the primary stakeholders and clients get what they want to know from the evaluations and accordingly, identify core questions for the evaluation.

7.6 Sri Lankan Case study on National Irrigation Rehabilitation Project (NIRP) selection and development of Key Evaluation Questions

Step 1—Review the Theory of Change and use it as a basis for the formulation of KEQs

Figure 23: Theory of Change—National Irrigation Rehabilitation Project



Step 2—KEQs are categorised based on the OECD/DAC evaluation criteria, namely: relevance, coherence, efficiency, effectiveness, impact and sustainability.

1. Relevance:
   - Is the project consistent with the National Irrigation subsector strategies and policies of the government?
   - Has the project included the needs and priorities of the farmers?
   - Are the project components and activities balanced and relevant to the Goals and Purpose?
   - Are the selected schemes, target groups and target areas appropriate?

2. Coherence
   - Is the intervention compatible with other interventions, sectors and institutions?
   - Did the project have any influence on other programmes, interventions and are there inter-connectedness, inter-linkages, synergies and trade-offs?

3. Efficiency
   - Were the schemes rehabilitated successfully including all civil works, especially Headwork and Delivery system (Channels)?
   - Does the cost incurred justify the degree of achievement of outputs and outcomes? (Comparison with similar projects)
   - Any alternative strategies that will achieve the same results?
   - What factors constrain or contribute to the efficiency of the project implementation process?

4. Effectiveness
   - Is the project strategy and sequence of activities appropriate?
   - Are the schemes functional? State of maintenance of the irrigation structures
   - Are the Farmer Organisations (FOs) ready to undertake O&M?
   - Do the FOs have the capability to generate funds to meet operations and maintenance (O&M)? How is the performance of the Fos?
   - What are the hindering or contributing factors for project effectiveness?

5. Impact
   - Has the overall goal been achieved?
   - Has the intervention institutionalised and improved water management practices?
   - Was there an increase in yield, cropping intensity and paddy production in the schemes?
   - Did the intervention improve the income and living standards of farmers in the schemes?
   - Any impacts on the institutions?
   - Any external factors which influenced the impacts positively or negatively?
   - Any unintended effects and impacts of the project?
   - Did the project have any influence on the policy?

6. Sustainability
   - Will the termination of the project affect the outcomes and impacts?
   - How far the schemes were successfully handed over to FOs for maintenance?
   - Have the FOs taken over O&M responsibility for minor schemes and selected F&D Cannels in major/medium schemes?
   - Are the FO's able to manage schemes, resolve water disputes, venture into marketing and input suppliers etc. and are the FOs economically and organisationally empowered through commercial activities?

CHAPTER 8

Step 6: Evaluation Design Matrix

8.1 Evaluation Design Matrix

The Evaluation Design Matrix is an essential tool for planning and organising an evaluation. It is simply a table with one row for each evaluation question and columns that address evaluation design issues, such as data collection methods, data sources, analysis method and reporting. The Evaluation Design Matrix is developed during inception once evaluation questions and data sources are agreed and it should be used by the evaluation team throughout the conducting of the evaluation.

It is the main analytical framework for evaluation and reflects key evaluation questions (KEQs) and sub questions to be answered and helps the team to consider the most appropriate and feasible method to collect data for answering each question.

Evaluation Design Matrix includes:

- KEQs to be answered;
- How judgement will be formed: Indicators on which the answer will be based should be specified; and
- Expected information sources and methods

The completed "Evaluation Design Matrix" is included in the inception report.

8.2 Steps involved in the development of the Evaluation Design Matrix

Step 1: Formulation of evaluation question under OECD/DAC Criteria: The purpose and intended use of evaluation with the 'Theory of Change' is used as a basis to formulate evaluation questions according to appropriate evaluation criteria.

Step 2: Appropriate indicator for judgement: Based on evaluation questions with sub questions, an appropriate indicator will be identified to judge the response to the questions. The indicator on which the answer will be based should be specified clearly.

Step 3: Data Collection Methods and Sources: The evaluation question will determine the information requirement which will be supported with data collection tools, methods and sources for each evaluation question.

Step 4: Analysis and Conclusion: Analysis will be conducted and conclusions and findings will be derived based on the indicator, data sources and methods.

Figure 24: Evaluation Design Matrix

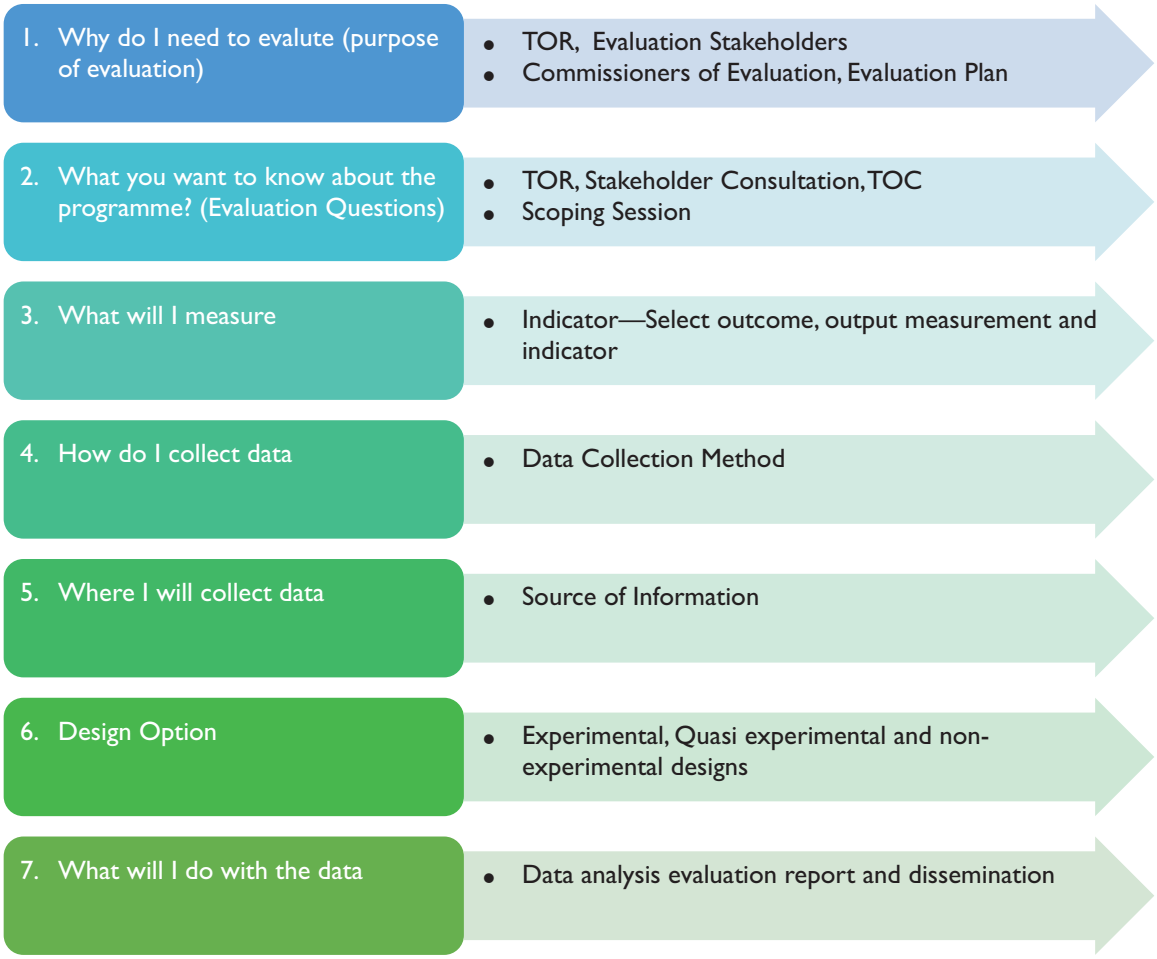
Evaluation Criteria / Evaluation Questions	Indicators	Data Source	Data Collection Methods

- A systematically documented Evaluation Design Matrix will form an important output in the inception report.
- The Evaluation Design Matrix should be a key product reviewed by the evaluation steering committee and evaluation reference group and approved by those committees as part of the inception report.
- The evaluation team will collect primary and secondary data to measure the indicators that have been identified, using the methods and tools agreed in the evaluation matrix.
- The evaluation questions can be answered by desk review, focus group discussions, key informant interviews, community interviews, observations, questionnaire survey covering quantitative, qualitative and mixed data and information.
- Triangulation is a validation method which uses different methods to answer the same question. Even if the evaluation question can be answered using only one method, it is often preferable to combine multiple methods to answer the same question in order to gain a more complete understanding of the issue, with more confidence in the findings. When evaluators use different methods to validate their findings, the process is called 'triangulation'.
- Based on the data and analysis, the findings and conclusion should be set out against the evaluation questions.
- The Evaluation Design Matrix should be included as an annex to the final evaluation report.
- The matrix will be used by the commissioner and steering/reference groups to review and check whether evidence has been collected, analysed and triangulated, as proposed in the matrix.

Figure 25: A typical World Food Programme Evaluation Design Matrix is given below:
(Source: World Food Programme)

Criteria: Relevance					
Evaluation Question 1: Was the intervention aligned to meet the food security and nutritional needs of the society?					
Sub questions	**Measure / Indicator of Progress**	**Source of Information**	**Data Collection Methods**	**Data Analysis Methods/ Triangulation**	**Evidence availability / reliability**
These questions are taken from TOR (explicitly or implicitly and further refined based on new information and discussion with stakeholders during the scoping session or inception phase. Example: Sub questions: Were the activity's targeting and transfer modalities appropriate to food security and nutrition needs of women, men, boys and girls from different groups?	The indicators or measures determine how performance or progress is judged for each sub question Example: perceptions regarding the degree to which needs of different groups were identified appropriately; and targeting was done based on needs % of beneficiaries who say that service met their needs (by group) The degree to which beneficiaries feel/perceive that the service was tailored to their needs. Evidence of different type of service being provided based on needs.	Sets out where the evaluation team will get information and data to answer each question. This stage is critical to informing the evaluation design Example: 2015 WFP VAM Analysis Government needs analysis study (2014) UNHCR evaluation 2015 Data from Key Informant Interviews Data from beneficiary focus groups (held separately with women and girl beneficiaries)	Covers what detailed data collecting methods will be used to collect the required data and information for each question. This can include quantitative beneficiary surveys; key informant interviews; desk review etc. Example: Document review using a structured framework Key Informant Interviews with: Co-operating partners Ministry representatives partner UN + donor representatives Focus groups with beneficiaries	Documents how all data that is collected is analysed to ensure they can answer the Evaluation questions. It helps to avoid collecting data that is not useful, and clearly shows how data is triangulated. This can include, regression analysis, statistics, qualitative analysis Example: Narrative/thematic analysis of secondary data Discourse analysis of primary data (interviews/focus groups) Data disaggregation (women/vulnerable groups)	Strength of evidence for each evaluation question. Can be recorded as colour coding (green/ amber/red); or numerically (3=strong, 2=fair, 1=weak), or with narrative descriptors (strong, fair, weak). Example: 3 (strong)

Figure 26: Detailed Steps in the Planning and Design of Evaluation



8.3 Data Collection Methods

The evaluation Design Matrix should include a clear description of what data will be collected, by what method and sources. The Evaluation Design Matrix will demonstrate how the data collection methods and sources are closely linked to the key evaluation questions. The following data collection methods are generally used in evaluation and are given reference in the Evaluation Design Matrix,

a) Desk Review or Document Review
b) Survey
c) Key Informant Interviews
d) Focus Group Discussions
e) Community Interviews
f) Direct Observations
g) Case Study

8.4 Evaluation Design Options

Evaluation Design is a structure created to produce an unbiased appraisal of a programme's benefits. Evaluation Design attempts to address two main concerns:

1. Measuring the expected results of an intervention
2. Attributing the results to the intervention and ensuring that it is not caused by some other factors.

There are three types of design options generally used in evaluation.

1. Experimental Designs
2. Quasi-experimental Designs
3. Non-experimental Designs

There is no perfect design. All evaluation design options have trade-offs on time, cost, certainty and practicality. Generally, the more certain you want to be about your programme's outcomes and impacts, the more costly the evaluation is. If the evaluation is intended to estimate impact that can be attributed to a specific intervention, the evaluation will include some kind of experimental or quasi-experimental designs, which typically include a control or comparison group. Experimental or quasi-experimental designs estimate the counterfactual by creating a Comparison Group or Control Group—a group who are like the people in the treatment group in all ways except in not receiving the programmatic intervention.

A complex evaluation design requires greater expertise and costs more but allows for greater confidence in its findings.

8.4.1 Experimental and Quasi-Experimental Designs

The Randomised Control Trial (RCT) is much like a scientific experiment and it is often considered the gold standard. One group (the "treatment group") experiences the intervention and the other group (the "control group") does not experience the treatment or intervention. The two groups are compared to see whether the intervention made any difference. The task is not only to show that the outcome occurred but to make the case that the outcome can be attributed to the intervention and not some other factors. In classical evaluation design, this problem of attribution is addressed by randomly assigning the potential pool of participants into treatment and control/comparison groups. The challenge is to create a comparable treatment and comparison group.

In other words, to know whether the intervention has made a difference, you need to know what would have happened if the same users did not receive the intervention (counterfactual). This enable you to attribute changes to the interventions, rather than other factors like motivation, another programme or family influences.

The counterfactual is a situation or condition which hypothetically may prevail for individuals, organisations or groups under, without project intervention scenario. It hypothetically develops a model as to what would have happened if the treated group did not receive the treatment or intervention. Experimental and Quasi-experimental research designs estimate the counterfactual by creating a comparison group or a control group – a group who are like the people who received a programme in all ways except in not receiving the programme. Counterfactual answers the question "What would have happened to the programme participants if they had not participated in the programme?"

The control group must either be selected completely at random or otherwise be very carefully selected to have very similar characteristics. Otherwise, you cannot be sure that any apparent differences in the results at the end are not the result of differences that were already there at the start and therefore nothing to do with your intervention.

The evaluation tries to establish whether the observed outcomes (project objectives) can be attributed to the project concerned. The research design enables the evaluator to establish the cause and effect relationship between the project activities and observed outcomes. For example, a nutritional project may be evaluated in terms of whether or not the project caused a reduction in the level of infant mortality.

Clearly these observed outcomes could have been caused not by the project which is being evaluated, but by other factors. The reduction in the level of infant mortality may have been caused by a post-natal programme working in the area rather than a nutritional project. These alternative explanations for an observed change can threaten the validity of the hypothesis.

To be able to attribute an observed change to a project activity, these possible alternative explanations for the change must be ruled out. The experimental research design enables researchers to establish the relationship between cause and effect by ruling out alternative explanations for change.

Attribution is defined as a causal link between observed or expected changes and a specific intervention. When the observed change in outcome is caused exclusively by the intervention and not by other factors, such causal change is defined as attribution. Changes are produced solely or wholly by the programme, the outcomes are fully attributed to the intervention. The attribution refers to both isolating and estimating accurately the particular contribution of an intervention and ensuring that the causality runs from the intervention to the outcome.

Experimental research designs or RCTs can be appropriate when it is possible to randomise people or sites. The use of research design in evaluation work is an area of great debate and argument.

The difficulties with RCT are:

- You need large samples. RCT is meaningful if there is a large control group with very similar characteristics to the users.
- RCT can be expensive. It is not a cost-effective design. RCT cannot tell you everything. RCT cannot tell you why something is effective or ineffective.
- Learning everything about the project on how it worked is not answered by RCT. Theory based evaluations do have these capabilities.

One of the primary purposes of the evaluation is to determine if the programme intervention has the desired effect. It is necessary to ensure that the change was a result of the intervention and not due to other factors or influences. A rigorous evaluation design helps to ensure this.

8.4.1.1 Experimental Designs

Typically, the most rigorous evaluation design is experimental design, where participants are randomly assigned to a programme or control group. It involves randomly assigning participants to a treatment or control group. The participants in the programme group receive the intervention, while those in the control group do not. Some form of pre/post assessment is provided to both groups and the results of these are compared to determine if there are differences between the groups. This method rules out alternative explanations and other influences.

This type of design is considered the "gold standard" against which other research designs are judged, as it offers powerful techniques for evaluating cause and effect.

8.4.1.2 Quasi-Experimental Designs

Participants are not randomly assigned to programme and control groups. it does not have a random assignment component, but may involve comparing a treatment group to a similar group that is not participating in the programme, to estimate the effect of the treatment or interventions. The challenge is to identify the comparison group.

8.4.1.3 Non-Experimental Methods

Non-experimental designs may include pre and post intervention studies with no control or comparison group. Non-experimental evaluation studies are likely to produce findings regarding programme outcomes. But they cannot control for extraneous factors that influence the outcomes.

Challenges and concerns in research designs should address the influence of alternative explanations or other factors.

Example: Child growth and eating carrots
If you take children's heights at the age of six (6), and feed them large amounts of a specific food for 3 years—say carrots, and measure them again at the end of the period, you will find that most of them are considerably taller at 9 years of age than they were at 6 years of age. You might conclude that eating carrots made the children taller, because your research design offered no basis for comparing children's growth. Hence, non-experimental designs experience challenges in attributing outcomes to the specific intervention as they cannot rule out alternative explanations and other influences.

Figure 27: Design Options

Research Design Methods	Features
Experimental Method	Experimental design is based on the selection of the comparison and treatment group through random sampling. It compares the treatment group to the comparison group to separately identify the change in outcome caused by the intervention. The control group participants are randomly assigned and the method rules out alternative explanations and other influences. It is a powerful technique for evaluating cause and effect and is considered a statistically rigorous evaluation design to identify attribution through intervention. It is necessary to ensure that the observed change in outcome or impact is the result of an intervention and not due to other factors or influences
Quasi-Experimental Designs	Quasi-experimental design is based on a comparison group that is chosen by the evaluator and is not based on random sampling. The method compares the treatment group to a comparison group—a similar group that is not participating in the programme, to estimate the effect of the intervention or treatment. Participants are not randomly assigned to the programme and control group. Less statistically rigorous whe compared to the experimental design
Non-Experimental Design	Non-experimental design may include pre and post intervention studies with no control or comparison group. Not a statistically rigorous method.

8.4.2 Validity of Research Designs

There are two kinds of threat to the validity of the research design

i. Threats to Internal validity—This is to examine whether the change produced is attributed to the intervention or due to other influencing factors (alternative explanation). If there is no way to separate their effect from those of the programmes, the observed change cannot be attributed to the intervention. These are called threats to internal validity.
Alternative explanations may include history (background, culture, experience, education etc.), Maturation (natural, physical, social, psychological process), measurement (scale) or selection of participants.

ii. Threats to External validity—This refers to the results. Are the results of the interventions likely to apply to other situations, contexts and people? External validity is about being able to apply your research results to other circumstances or context. It is all about "generalisability". There is a threat in applying the results to the other context, circumstances and situations.
The design matrix is a visual way to map your evaluation plan with clear evaluation questions under the evaluation criteria. It includes deciding on the questions you will answer, choosing an appropriate design for the question and situation, determining how you will construct the measures (KPIs), locating likely sources of information and data, deciding on the data collection strategy, sampling, design options and likely data analysis techniques. The outline of the evaluation design matrix is set out below:

Figure 28: Evaluation Design Matrix

Evaluation Criteria / Evaluation Questions	Indicators	Data Source	Data Collection Methods

8.5 Rating and Scoring

After developing the Evaluation Design Matrix, it will be submitted at the scoping session and consulted and agreed upon by the stakeholders. Once it is agreed on by the steering committee, reference group and peer review group, the Evaluation Design Matrix should be incorporated into the inception report. The Evaluation Design Matrix will set out the evaluation plan and will be used as a planning and design document in the conducting of the evaluation. During the conducting of the evaluation, the KEQs will be answered and supported by data collection and analysed into indicators to assess the performance of the project, in terms of the evaluation questions and relevant criteria.

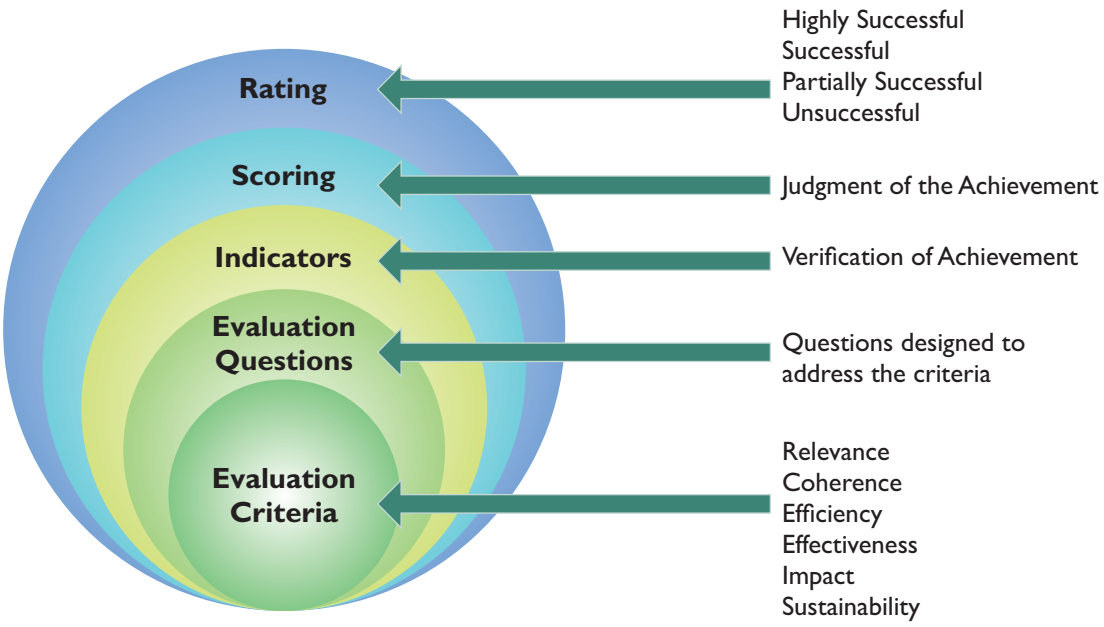
As a next step, it is necessary to rate the project into a four-band rating criteria as given below:

1. Highly Successful (HS)
2. Successful (S)
3. Partly Successful (PS) and
4. Unsuccessful (US)

Based on the DAC Evaluation Criteria and considering the evaluation questions and indicators, a scoring is assigned, based on the analysis of the responses and indicators. The Asian Development Bank and some donors and partner countries use the scoring method for rating the performance of projects as highly successful, successful, partly successful and unsuccessful. The rating methodology is given below:

Figure 29: Rating System

Criterion	Weight	Rating Description	Rating Value
1. Relevance	20%	**Highly Relevant** **Relevant** **Partly Relevant** **Irrelevant**	**3** **2** **1** **0**
2. Efficacy	25%	Highly Efficacious Efficacious Less Efficacious Inefficacious	3 2 1 0
3. Efficiency	20%	Highly Efficient Efficient Less Efficient Inefficient	3 2 1 0
4. Sustainability	20%	Most likely Likely Less Likely Unlikely	3 2 1 0
5. Institutional Development and Other Impact	15%	Substantial Significant Moderate Negligible	3 2 1 0
Overall Assessment (Weighted average of A1, A2, A3, B and C)	Highly successful (HS): Overall weighted average (OWA) is > 2.5 and none of the 5 criteria has a score of less than 2; Successful (S): OWA is between 1.6S 2.5 and none of the 5 criteria has a score of less than 1; Partly Successful (PS): OWA is between 0.6PS 1.6 and number of criteria receiving a rating of less than 1 should not exceed 2; Unsuccessful (US) : OWA is < 0.6		



8.6 Evaluation Methodology: Multiple Methods

In evaluation there is no methodological mono-culture. There is no one solution or methodology. The way forward is not rigid adherence to any one methodology, but openness to diversity, innovation and change. The general practice has been to adapt methodological pluralism. In the early 1960s and 1970s, the evaluations were highly based on the economic analysis/appraisal of projects. Therefore, financial and economic analysis based on cost benefit analysis techniques were extensively used for assessment of the performance of projects.

Subsequently, many bilateral donors such as German, USAID and a number of European donors including the UK started using Logical Framework Analysis (LFA) with Key performance indicators as a methodology for evaluation, to complement and supplement the financial and economic cost-benefit analysis.

Subsequently in late 1980s, Robert Chambers and a group of Evaluation Academics emphasised participatory approaches with the active involvement and engagement of stakeholders. In the early 1990s, the OECD initiated evaluation criteria focusing on relevance, efficiency, effectiveness, impact and sustainability, with evaluation questions and methodologies to assess the performance of development interventions. In the early 1990s, in addition to financial and economic perspectives, social, environmental, institutional, gender and sustainability aspects received much attention in the evaluation methodology. In 2015, with the 2030 development agenda, the evaluation methodology was expanded to align with the SDGs and included cross-cutting thematic areas such as climate change, the environment, human rights, peace building, gender and partnerships. Thus, the evaluation methodology has undergone rapid changes in focus, tools, the techniques used and approaches adapted.

In the current context, evaluation uses multiple methods which include quantitative, qualitative and mixed methods. The 'theory-based approaches' are gaining more importance.

There are two methodologies for evaluation:

1. Rapid Appraisal/Evaluation Methodology
2. Rigorous Appraisal/Evaluation Methodology

8.6.1 Rapid Appraisal/Evaluation Methodology

Under this method, evaluations are conducted in the old style of a 2 to 3-week mission to evaluate the project or intervention, using quick methods. Generally, such evaluation missions are sent from donor organisations and though well intentioned, the hardworking mission uses rapid methods to support evaluation findings. The mission generally uses "quick and dirty methods" which are not grounded and supported with in-depth investigations, analysis and assessments. They generally used desk reviews, literature reviews, key informant interviews, field visits and observations, focus group discussions, community interviews, rapid semi-structured surveys etc. This methodology has been criticised for its limitations as it fails to rigorously cover complex and diverse issues and context specific issues. Rapid Method also undertakes 'wind shield surveys' and observations and is critically seen as "development tourism". In terms of time and cost, these methods have advantages. However, in terms of quality, accuracy and depth and breadth of the coverage, the method is considered as weak. Therefore, it is considered "quick and dirty methodology".

8.6.2 Rigorous Appraisal/Evaluation Methodology

This takes an average time of 8 to 12 months and collects extensive relevant data and information, using experimental and quasi-experimental methodologies with rigorous statistical methods to attribute the observed changes to the intervention under cause and effect principles. The rigorous methodologies involve extensive engagement of stakeholders and use financial and economic cost benefit analysis and

cost-effectiveness analyses, including social and environmental analyses to evaluate the intervention. Rigorous methodologies examine the technical, marketing, financial, economic, social, environmental, institutional, gender, policy and political and sustainability analyses to ensure the analyses are rigorous in nature. Before and after analysis, with and without project analysis, LFA analysis, cost-effectiveness analysis, outcome-mapping techniques and participatory rural appraisal and rapid rural appraisal methods are used with quantitative, qualitative and mixed methods.

Therefore, the rigorous nature of the assessment with case studies and counterfactual analysis provides a more accurate assessment of the results. However, although the accuracy and rigorousness are high, the costs and amount of time taken are comparatively high.

8.6.3 **Rapid Rural Appraisal** (RRA)

Rapid Rural Appraisal (RRA) is a social science approach that emerged in the early 1970s, to quickly collect, analyse and evaluate information on rural conditions and local knowledge, in close cooperation with the local population in rural areas. In RRA, information is extracted from the rural community by outsiders, through quick methods and approaches. RRA techniques are an attractive alternative to conventional surveys. RRA is an efficient and cost-effective way of learning by outsiders, particularly about rural needs, conditions and projects by having discussions with a rural community. RRA includes semi-structured interviews, focal group discussions, community interviews, key informant interviews, field visits and interviews and is considered as a way of extracting information on projects and programmes from rural community using these methods.

8.6.4 **Participatory Rural Appraisal** (PRA)

Participatory Rural Appraisal (PRA) is an approach used by evaluators to obtain the knowledge and opinions of rural people about the planning, management and evaluation of programmes and projects, using voiceless methods. Ordinary illiterate people would not feel comfortable to communicate in scientific terms. Robert Chamber argued that the PRA enables rural people to analyses their own reality. Those who had voices dominated the discussions and feedback in RRA. Therefore, PRA uses voiceless visual methods such as village mapping, diagnostic walk, transect walk, venn diagram, wealth ranking, scoring and ranking and other related methods which do not involve voice-based expressions. Information is owned and shared by local people, while the outsiders only facilitate the activity. Therefore, this method of data and information collection is considered to be empowerment-oriented and owned by the rural community.

8.6.5 **Financial and Economic Analysis**

Financial and Economic analysis is a critical component of a comprehensive project/ programme evaluation methodology. It is based on the cost and benefit streams of a project, over a multi-year timeframe. Cost benefit analysis was used extensively in the 1960s and 1970s. Both financial and economic appraisal/analysis predicts the profitability of a project by comparing the project's expected benefit streams and cost streams over a multi-year period. The financial appraisal analyses the profitability of a project from a private businessmans' or commercial entities' perspective and calculates this by using the market price or actual price paid for inputs (labour, construction and fuel costs) and the actual price received for the project outputs.

Eliminate transfer payment and adjust the market price to the economic price

The economic appraisal/analysis, the benefit stream and cost streams of a project are identified from the perspective of a national economy or society. The cost stream includes capital costs and operating costs. The price used in economic analysis is not the market price but the opportunity-cost based price which is

known as the shadow price. Shadow price is an adjusted market price which reflect the opportunity-cost based price of the society. The market price is distorted due to taxes and subsidies. Hence, the taxes, duties, subsidies are transfer payments that need to be removed from the market prices to arrive at economic prices/shadow prices. Similarly, when there is high unemployment of 20% in the area, the market may not reflect the real wages. This will push people to work for a lower-than-market wage. Hence, a conversion factor (CF) of 0.8 has to be applied to get the real opportunity-cost based wage. This is called shadow wage rate. This will be included into the economic cost benefit analysis. Conversion factors are used to convert market price into economic prices/shadow prices, for the purpose of undertaking an economic analysis.

Quantifying externalities

The project may incur social costs which will not be taken into consideration by the private businessman in the financial analysis. On the other hand, from the national and social perspective, the economic analysis should include such social and environment costs. For example, an agricultural processing plant may pollute a nearby river. An irrigation project may increase the incidence of water borne diseases or destroy the habitat of a particular species of wildlife. These social costs need to be included into the economic analysis, from a national perspective.

Traded goods and non-traded goods

The economic analysis considers the traded goods priced in terms of international border prices. When it comes to non-traded goods such as unskilled labour, the domestic market price has to be converted into an international equivalent price.

With and without project situation

The economic cost-benefit analysis should generally be undertaken with and without project analysis to identify the incremental benefits of the intervention. Similarly, when it comes to the cost, the "sunk cost" which is the cost already incurred, should not be taken into consideration.

Discounting permits comparison of costs and benefits

The cash flow occurring at different years requires the adoption of an appropriate discount rate to calculate the present value. Discounting is necessary because costs and benefits take place in different time periods and are therefore not directly comparable. Thus, discounting helps to bring the future cost and benefits into present value, which can be compared and aggregated.

Net PresentValue (NPV), Internal Rate of Return (IRR) and Benefit Cost (B/C) Ratio

The cost benefit analysis provides three main indicators. The NPV provides the net wealth generated and if it is positive (NPV>0), the project is evaluated as a successful project whereas if the NPV is negative (NPV<0), the project is rated as unsuccessful, leading to a depletion in the net wealth.

The Internal Rate of Return (IRR) is the discount rate at which the net present value is zero. IRR provides the rate of return of the project. If IRR>Cost of Capital (COC), the project is successful, whereas if the IRR<COC, the project is unsuccessful. The IRR calculated in the financial analysis is known as the financial internal rate of return (FIRR) and it shows the commercial viability of the project from an investor or businessman's perspective. The IRR calculated in the economic analysis is called the "Economic Internal Rate of Return" (EIRR) and it shows the economic viability of the project, from the national or social perspective.

The Benefit Cost Ratio (B/C ratio) is the present value of benefit to the present value of costs. If it exceeds (BCR>1), the project is successful, whereas if the BCR<1, the project is unsuccessful.

A sensitivity analysis which is known as the "what if analysis", is widely used in the cost benefit analysis under various risk and uncertainty conditions and to analyse different scenarios, such as best case, medium case and worst case scenarios.

The cost-effective analysis evaluates which project design option provides the desired results at the lowest cost. This is also used as a tool for evaluation.

8.6.6 Logical Framework Analysis (LFA)

The LFA is used as a powerful planning, management, monitoring and evaluation tool. It demonstrates the Theory of Change with underlying assumptions and uses KPIs to measure performance at different levels, with data collection methods and sources. The structure of the LFA consists of a 4 x 4 matrix in which the rows represent the level of project objectives, including the means required to achieve them (Vertical Logic), while the columns indicate how the achievement of these objectives can be verified (the horizontal logic).

Figure-31: The Project Planning and Design Matrix (LFA Matrix)

Hierarchy of Objectives (HOO)	Objectively Verifiable Indicators (OVI)	Means of Verification (MOV)	Assumptions
Goal/Impacts	Measuring the achievement of Goals		
Purpose/ Outcomes	Measuring the achievement of Purpose		
Outputs	Measuring the achievement of Outputs		
Inputs	Summary of the project inputs: Costs and Resources		

The LFA matrix can be analysed from two perspectives. From a planning and design perspective—the Vertical Logic and from a management, monitoring and evaluation perspective—the Horizontal logic. The vertical logic represents the 1st and 4th column and it is used for planning and designing the project. The 2nd and 3rd column represent the horizontal logic and it is mainly used for monitoring the performance through objectively verifiable indicators and the means of verification which set out the source of collecting data. A model LFA developed for Inland fisheries project is set out below:

Figure 32: Model LFA—Inland fisheries Project

	Hierarchy of Objectives	OVI	MOV	Assumptions
Goals	• Increased nutritional status of the community with improved protein consumption • Supplementary income to farmers and fishermen	• Incremental protein consumption • Increased income of farmers by 30% and fishermen by 35%	DHS HIES	• Operation and maintenance arrangement in place (financial and institutional) • State patronage for inland fish production
Purpose	• Increased fingerling production • Increased inland fish production	• Increased fingerling production per annum • Increased Inland fish production per annum	Field Survey Reports of FDCS	• Consumer preference for inland fish production (odour and taste) • Price stability
Outputs	• Upgraded Seed Production Centre with hatcheries, rearing ponds, brood stock ponds that can produce 120mn fingerlings per annum • Upgraded 1000 Seasonal Tanks • 5000 backyard ponds	• 6 upgraded hatcheries • 20 hectares of nurseries developed • 15 hectares of rearing ponds developed. • 25 hectares of brood stock ponds developed • 1000 seasonal tanks developed • 5000 backyard ponds	Project Completion Report (PCR)	• No conflict of interest between farmers and fishermen • Favorable climate • Religious priests are supportive of pond culture • Appropriate technology • Disease control mechanism in place.
Inputs	• 20 Mn US$ • 1000 existing seasonal tank • Existing Seed Production Centre • Technically Qualified Staff	Investment	Financial reports of PMU Progress Reports	• Contractors are capable • Quality technology procured • Counterpart fundss/ staff are available

8.6.7 The Methodological Debate

The economic appraisal method that uses the cost benefit analysis has come under increased criticism. Similarly, the experimental research approach in evaluation work has also been increasingly criticised in the evaluation literatures. The appropriateness of the particular methodology depends to a great extent on the context within which the project operates. Therefore, based on the context, a multiple method with methodological pluralism is a more appropriate way forward.

8.7 Sri Lanka Case Study Application: Aquaculture Development Project Planning and Design of Evaluation:

Fish is the principal source of protein for the country's population. The annual per capita fish consumption is estimated at 15–16 kilograms (kg), in the rural areas and this is lower than the 21 kg recommended by the Sri Lanka Medical Research Institute. The Asian Development Bank (ADB)-funded Aquaculture Development Project was aimed at increasing inland fish production to support the nutritional requirements of the rural community who did not have access to sea fish. As inland fish were to be reared in the irrigation schemes, the project also had the subsidiary objective of delivering a supplementary income to the farming community.

The cost of the Project was US$20 Mn and the project was planned to be implemented in 4 years, commencing in 1990. The project was to upgrade 6 inland fishing hatcheries at Inginiyagala, Muruthawela, Udawalawe, Polonnaruwa, Dambulla and Ginigathena as major seed production centres, to produce an incremental fingerlings production of 120 Million per annum

In addition to strengthening the hatcheries, the project was also planned to support the development of Nurseries and rearing ponds with 20 hectares and 15 hectares respectively. Moreover, 25 hectares of Brood stock ponds were to be constructed. The project targeted the development of inland fisheries in 1000 existing seasonal tanks covering 10,000 hectares, to produce 7,000 metric tons of incremental inland fish production. The project also had a small component to develop pond culture as a backyard activity, at the household level. Accordingly, 5,000 households were targeted and ponds were to be developed by the households and fingerlings were to be provided by the Aquaculture Centres. A group of priests objected to this practice and this appeared to be a risk factor in this project.

The economical internal rate of return was estimated to be 20% and the financial internal rate of return was estimated to be 10%. The project intervention was expected to increase the income of the farmers by at least 30%. ADB has appointed you to serve as a member of the evaluation team, to evaluate the Aquaculture Development Project and the Head of the Evaluation Mission has requested you to assist in the planning and design of the ex-post evaluation.

(a) Establish the "Theory of Change" for this project.
(b) Formulate suitable Evaluation Questions to assess the performance of the project at an ex-post situation.
(c) Formulate an Evaluation Design Matrix to map out your evaluation plan, with clear evaluation questions (select one or two KEQs) under each criteria and demonstrate it in the Evaluation Design Matrix, as set out below:

Evaluation Criteria/ KEQs	Indicators	Data Source	Data Collection Method

Outline Answers

(a) Theory of change

Inputs	Outputs	Purpose	Goals
20 Mn US$ Existing Seed Production Centres 1000 existing seasonal tanks Technically qualified staff staff	Upgraded Seed Production Centres with hatcheries, breeding facilities, rearing ponds and brood stock ponds that can produce 120 mn fingerlings per annum Upgraded 1000 Seasonal Tanks 5000 backyard ponds at household level	Increased fingerling production Increased inland fish production	Increased the nutritional status of the community with improved protein consumption Learning achievement of children Supplementary income to farmers and fishermen

(b) Evaluation Question under OECD/DAC Evaluation Criteria

Relevance:

1. Is the ADB Supported Inland Fisheries Project relevant and aligned to the Fisheries Sector National Development Plan and the strategy and policy of the Government of Sri Lanka and does it meet the needs, priorities and challenges of the target group in the project area?

2. Is the intervention logic of the project articulated through the project strategies, and are the components and activities consistent and relevant to achieve the project purpose and objectives? Are any irrelevant components included or any important components missing?

3. Are the project components and activities balanced in terms of the social and cultural contexts?

4. Has the project selected appropriate, promising and relevant inland fishing centres and seasonal tanks and implementing modalities?

Coherence

1. Is the intervention compatible with other interventions, sectors, institutions and programmes?

2. Does the intervention have inter-connectedness, inter-linkages, synergies and trade-offs within the sector or any other sectors of the economy?

Efficiency

1. Was the rehabilitation of 6 inland fisheries centres and 1000 seasonal tanks of 10,000 hectares and the establishment of 5,000 ponds at the household level completed successfully, as per the SAR target, within the time and cost and meeting the quality standards?

2. Were the activities carried out in a cost-efficient manner (least cost) and did the cost incurred justify the achievement of output and outcomes?

3. Did the project examine any alternative strategies and implementing modalities to achieve the expected results?

4. Were any serious issues faced by the project during implementation? Was the project coordinated, managed, monitored and evaluated efficiently?

Effectiveness

1. Has the ADB supported project produced or is it likely to produce the desired outcomes (increased fingerling production of 120 mn; inland fish production of 7000 MT)? Has it increased the per capita fish consumption from 15 kg to 21 kg?
2. Has the project created any unintended outcomes?

3. Which centres are functioning well and why? Which centres are not functioning well and why?

4. What factors have been instrumental in the non-achievement of the envisaged outcomes as set out in (1) above?

5. Is there a continuous demand for inland fish and any positive and negative aspects? (taste, smell, variety, dry-fish etc.)

6. Harvesting, storage, marketing and distribution arrangements and issues if any?

Impacts

1. Has the project achieved or is it likely to achieve its long-term objectives? i.e, improved consumption of fish, improvement of the nutritional status, supplementary income to the farmers. What are the unintended impacts if any?

2. Has the project created or is it likely to create any positive or negative/direct or indirect/intended or unintended, economic, social or environmental impacts?

Sustainability

1. Were the centres successful in generating funds to meet the O&M cost by charging a reasonable price for the sale of fingerlings?

2. Is the relationship between Farmer Organisations and the Fisheries Community smooth?

3. Are the project benefits likely to continue after donor funding is terminated? How will the project be actively up-scaled?

4. Major factors influencing the achievement of sustainability of the programme?

Evaluation Design Matrix (outline Answer)

Evaluation Questions	Indicators	Data source	Data Collection Methods
Relevance			
1. Is the ADB Supported Inland Fisheries Project relevant and aligned to the Fisheries Sector National Development Plan and strategy and Policy of the Government of Sri Lanka and meet the needs, priorities and challenges of the target group in the project area?	• Degree of alignment of project strategies and objectives to National Development Framework and Fisheries Sector Plan (%) • % of beneficiaries who said that the project met their needs	• National Development Plan • Fisheries Sector Plan • Project Proposal and Appraisal Report including LFA • Interim Evaluation Report • Perception of beneficiaries	• Desk Review (Literature Review) • Key Informant Interview • Policy Level Stakeholder Consultation • Focus Group Discussion
2. Is the intervention logic of the project articulated through the project strategies, components and activities are consistant and relevant to achieve the project purpose and objectives ? Any irrelevant components included or any important components missing ?	**Perception of stakeholders on** • Well articulated intervention logic or theory of change Perception of beneficiaries and stakeholders • All important problems being addressed	• Project Proposal and Appraisal Report with clearly articulated LFA • Interim Evaluation Report • Beneficiary Consultation Feedback Report	• Desk Review • Beneficiary consultation • Policy Level Stakeholder Consultation
Efficiency			
1. Were the rehabilitation of 6 inland fisheries centres and 1000 seasonal tanks of 10,000 hectares and establishment of 5,000 ponds at the household level completed successfully as per SAR target, within the time and cost meeting the quality standards?	• Timely completion of project deliverables • Process Indicators – time overrun cost overrun and quality compliance	• Project Completion Report • Interim / Terminal evaluation Report • Project Management Information System (Database) • Beneficiary Consultation / Operational staff of the Centres Feedback Report • M&E Reports	• Desk Review • Key Informant Interview • Focus Group Discussion • Field Observations / Field Visits • M&E Reports/ Beneficiary Feedback Report

Evaluation Questions	Indicators	Data source	Data Collection Methods
Effectiveness			
1. Has the ADB supported project produced or likely to produce the desired outcomes ? – increased fingerling production of 120 mn; Inland fish production of 7000 MT and increased the per capita fish consumption from 15 kg to 21 kg	• Project achievement on fingerlings production, inland fish production and per capita fish consumption	• Performance Report of the Inland Fisheries Centres • Fisheries Extension Officers Report and Fisheries Societies Report • Consumer Feedback Reports and surveys	• Desk Review (Literature Review)/ reports • Key Informant Interview • Focus Group Interviews and Community interviews • Rapid consumer survey • Feedback from execution officers (focus group discussion)
Impact			
1. Has the project achieved or likely to achieve its long-term objectives? ie, improved consumption of fish, improvement of the nutritional status, supplementary income to the farmers, unintended impacts if any	• Achievement of targeted outcomes ? per capital fish consumption increased from 15kg to 21kg • Increased supplementary income to farmers • Unintended Impacts	• Field Survey Report • Health / Nutrition report • Field Discussion Feedback	• Desk Review / nutrition report and survey (Literature Review) • Key Informant Interview • Focus Group Interviews and Community interviews • Field Observation / Field visits
Sustainability			
1. Whether the project benefits are likely to continue after donor funding is terminated How ? will the project actively be up-scaled ?	• Cost recovery, self financing arrangement and up-scaling of services from centres • % of cost recovery • Demand for inland fish	• Revenue Statement of the Inland Fisheries Centres and Operational statement of the Centres • Consumer Demand survey	• Desk Review/ review Statements • Key Informant Interview • Focus Group Interviews and Community Interviews

CHAPTER 9

Step 7: Management and Conduct of Evaluation

The Management of Evaluation is structured into 4 areas to provide guidance and direction to ensure quality, conduct meaningful evaluations and promote the use of evaluation.

Management structure for Evaluation:

1. Evaluation Steering Committee
2. Evaluation Management Group
3. Evaluation Reference Group
4. Evaluation Peer Review

9.1 Evaluation Steering Committee

A Coordination Mechanism is needed to oversee the evaluation and this is called a steering committee which takes responsibility for ensuring the quality of the evaluation and approving the evaluation report.

The steering committee is supported by a management group to liaise with the service provider on a day-to day basis. The steering committee oversees and takes decisions on the overall evaluation. The TOR of the steering committee may include reviewing and approving the Evaluation TOR/Work Plan, appraising the peer reviews, reviewing the inception report, directing the incorporation of peer review comments, recommending changes and appraising the inception report, reviewing the draft final report and providing feedback on the draft report, making recommendations on the Management Response and Improvement Plan and working on the dissemination strategy.

The sponsors and funders of the evaluation, representatives from the executing and implementing partner organisations are all expected to provide broad direction and guidance for the evaluation and to review progress on the deliverables.

9.2 Evaluation Management Group

The Management Group comprises of evaluation managers from different agencies and partners and it is tasked with the day-to-day management of the evaluation under the guidance and direction of the steering committee.

9.3 Evaluation Reference Group

The Reference Group comprises of professionals and experts including technically competent officials who are entrusted with the responsibility of reviewing the TOR and the inception evaluation report and providing advice during the evaluation processes. The Evaluation Reference Group should advise on independence, credibility, utility and other principles including ensuring that standards, ethics and norms are complied with and practiced.

The Evaluation Reference Group will ensure the buy-in to the evaluation process by different stakeholders (policy-makers, donor, partner countries/CSOs) and give professional guidance for the evaluation.

The Evaluation Reference Group will endorse the evaluation TOR, comment on the TOR, comment on the evaluation report, ensure the dissemination strategy and the buy-in to the evaluation by stakeholders, to ensure its effective use.

9.4 Evaluation Peer Reviewer

Peer Review is an internationally accepted approach to strengthen the quality of evaluation, promote good quality evaluations and to ensure the maintenance of standards, credibility and quality assurance in evaluation.

Quality assurance takes place at different stages of an evaluation. The peer review examines the TOR, contents covered, appropriateness of key evaluation questions, adequacy of the literature review, robustness of the designs, whether the TOC and the proposed project design matrix/analytical frameworks are adequate. It also comments on data collection methods and sampling arrangements and on the quality of the field work, the stakeholder consultation and the draft report and its contents.

The Peer Reviewer should possess sector knowledge, expertise in methodology and knowledge of evaluation. The peer reviewer is recommended by the custodian ministry and approved by the evaluation steering committee and a standard honorarium is usually paid.

9.5 Data Collection

An evaluation planning or evaluation design matrix should include a clear description of what data will be collected, from which sources and when, as well as how the data will be analysed. Planning this information in the evaluation design matrix helps to demonstrate that the data collection and analysis is linked to the key evaluation questions (KEQs).

Collecting data is a major part of any evaluation but it is necessary to keep in mind that the data collection method follows the purpose of evaluation and the key evaluation questions. However, it is necessary to ensure that the data collected is authentic, credible and trustworthy. It is also necessary to ensure confidentiality, to respect the rights of beneficiaries and the right to privacy and to respect the local culture. Evaluation should ensure that evaluation ethics are strictly followed and adhered to. More insightful information can be gained if the data is disaggregated by gender, age, ethnicity and level of income. The data collection strategy can be categorised as follows

1. Primary and secondary data sources
2. Quantitative, Qualitative and Mixed Methods

Primary data is collected by the evaluation team especially for the evaluation study and it is usually collected by the evaluator through surveys and questionnaires. Secondary data is the data collected and recorded by another institution for a different purpose than the current evaluation. Secondary data is the existing available data collected from Census data, country level indices, World Bank data and other administrative data available at institutions and sub-national level agencies.

Data Collection Methods

Following are some of the data collection methods used extensively in evaluation.

1. Document Review/Desk Review—the Evaluator can collect information by reviewing existing documents printed material, reports and available literature. This may include reviews of project appraisal documents (PAD), mid-term reviews, Donor Mission Reports, media comments and project progress reports. These documents will provide insightful information to understand the project, its setting, progress achieved and issues and challenges.

2. Surveys—surveys are useful data collection methods for large sized data where standardisation is considered important. It may include close-ended questions where respondents are asked to select from a range of pre-determined answers. Close-ended questions are easy to analyse and generalise. On the other hand, open-ended surveys allow free flowing responses and while they may not be easy to tabulate and analyse, they provide in-depth and insightful information.

   In today's context, the survey method no longer uses a paper and pencil approach—that has been replaced by technology, where computer-assisted surveys, e-mail responses, web-based on-line-data etc., enrich the survey method. The survey method not only captures quantitative data but also covers qualitative data such as satisfaction, perception and attitudes.

3. Key informant Interviews
   Information is collected by talking and listening to key informants and may use structured or conversational methods with free-flowing interviews. Key informants are champions in a community, or persons with unique skills or professional backgrounds, with first-hand knowledge on the topic of interest, geographical setting or community.

4. Focus Group Discussions
   Focus Group discussions are usually conducted with 8-12 persons on a thematic topic. The discussions are in-depth and they capitalise on group interaction to generate data, information and insights. The interaction with the respondent stimulates rich information and insights and these discussions are generally conducted by a skilled moderator, to ensure all voices are heard.

5. Community Group Interviews
   These are interviews conducted with the community members in an open setting where perceptions, attitudes and insightful information to understand the planning, implementation and post-implementation issues can be discussed. This may include both the beneficiaries and non-beneficiaries of the project, thus helping to understand the comparative picture, with grievances and issues being discussed.

6. Direct Observations
   The Evaluator visits the project site and observes the working of the project and the project-setting, including the degree to which the beneficiaries are involved, through direct observation. Data and information is collected based on what is seen, heard or touched by the evaluator. It also helps the evaluator to understand things which were not expressed by the beneficiaries.

7. Case Study
   This is an in-depth examination of a particular case, a programme site or location. The case study is a study based on a small number of "typical examples". The case study relies on multiple sources of information and provides a complete picture of the selected project case study. It will provide rich information, analysis and exploration of the selected case study project area in the setting.

Factors considered in choosing the data collection method

The following factors are considered in selecting the data collection method in the planning, design and conduct of the evaluation:

1. Appropriate method to answer the KEQs.
2. Information should be authentic, reliable, credible and trustworthy.
3. The availability of resources in terms of time, funds and travel.
4. Potential to triangulate through multiple methods, to enable cross-checking and to increase validity.

5. The beneficiary's perspective is more important than the evaluator's perspective. Structured questionnaires and surveys are beneficial and easy to analyse from an investigator or evaluator perspective. But from the beneficiary's perspective, open-ended questions with loosely structured interviews are well received.

No single approach is perfect. A combination of quantitative, qualitative and mixed approaches with diverse methods is more suitable and appropriate.

9.6 Possible biases to be avoided in undertaking evaluations

Although evaluations are independent, objective, impartial and credible, it is important to ensure that checks and balances are in place, to avoid any biases in the design and conduct of the evaluation. The following are some of the biases identified in a conventional evaluation process.

- Spatial Biases—In undertaking evaluations, it is necessary to ensure peripheries are not avoided. There is a general tendency to avoid interior backward areas and peripheries in setting up the field mission.

- Project Biases—Failure to look into non-project areas. It is noted that in some evaluations, the evaluation team fails to understand the outcome of the intervention due to a lack of knowledge of non-project areas.

- Person's Biases—In some cases the evaluation team, during their field mission, visit and talk to knowledgeable key informants and elite groups who are educated, whereas the vulnerable highly marginalised stakeholders' view are not adequately captured. The field mission should not function as a rural development tourist but should see bad conditions and talk to the highly vulnerable, voiceless beneficiaries to get their perceptions and perspectives.
- Diplomatic Biases—Failure to see bad conditions based on an itinerary prepared in advance.

- Professional biases—There is always a perception that "experts know best". Moreover, sometimes indigenous technical knowledge (ITK) is not fully tapped. Outsiders fail to perceive the true problems (Robert Chambers).

In an evaluation, it is necessary to ensure these biases are avoided, to ensure quality and meaningful evaluation.

10 CHAPTER

Step 8 and 9: Evaluation Report, Dissemination and Management Response

The primary objective of the evaluation report is to inform the management of the results of the evaluation. The Evaluation report should articulate a comprehensive response to the expectations set out in the TOR. The management should ensure that the final report meets (a) the contractual obligation; (b) complies with professional standards, norms and ethics; (c) fairly and accurately assesses the performance and (d) delivers useful and credible responses to the KEQs

The presentation of evaluation results should follow a logical process following a credible method linking the KEQs to information collected and the conclusions and judgements given.

The following general quality criteria need to be satisfied:

1. Clear statement of the Evaluation Question.
2. Clear presentation of indicators and measures for assessment. The basis for value judgement made in the report should be explicitly stated.
3. A transparent account of the data collection source and research method is important.
4. Justification of Conclusions: Supporting evidences should be clearly presented and alternative explanations of findings should be explicitly considered or eliminated.
5. Impartial reporting: The perspective of all major stakeholders should be impartially reported in a balanced manner.
6. Clear statement of limitations: The limitations with regard to scope, coverage and depth of analysis should be explicitly stated.

10.1 Preparation of Evaluation Report

10.1.1 Inception Report

Only complex and high budget evaluations may require an Inception Report. The purpose of the inception report is to ensure that the evaluator understands the TOR. The Inception Report will contain the following information:

1. Adherence to TOR: The inception report should describe the conceptual framework planned and should address the criteria set out by the OECD on relevance, coherence, efficiency, effectiveness, impact and sustainability. It should provide the key evaluation questions (KEQs) with 2 or 3 specific evaluation questions, with data collection methods and sources.

2. Evaluation methodology: The Consultant should describe the chosen data collection method, data sources, sampling and indicators that will support the evaluation questions. (Figure 33)

Figure 33: Evaluation Design Matrix

Evaluation Criteria/KEQs	Indicator	Data collection method	Source of Data	Data Analysis

Site visits proposed should be indicated and planned meetings, discussions and workshops should be documented.

3. Work Plan: Inception report must contain the work plan which should indicate phases, timing, key deliverables and milestones, to keep track of the evaluation process.

Figure 34: Work Plan

Phases	Responsibility	Task	Proposed timeline	No.of man days
Inception/ Preparation	Evaluator	Study the TOR	1st to 10th July 2020	6 days
		Desk Review—review of project documents, TOC, sectors, interventions and business model; Key Evaluation Questions (KEQs)		
		Kick-off meeting with project staff		
		Develop EA methodology and KEQs		
		Inception report with work plan		
	Evaluation Management and Evaluator	Comments on inception report and finalisation	10th -14th July 2020	3 days
Consultancy /Data Collection and field information	Evaluator	Meetings / discussions with project staff including telephone and online meetings	14th to 25th July 2020	11 days
		Meetings and discussions with beneficiaries including producer cooperatives/local farmers/ producers and private sector partners, women and PWDs and local community leaders		
		Discussion with senior Divisional, District and Provincial and Policy Makers and Development Partner including the ILO		
		Desk Review on business models, M&E reports and other livelihood related policy, programme documents, discussions with senior officials, members of the PTF and SDG related stakeholders		
Data and Information Analysis	Evaluator	Prioritisation of KEQs; improvements to the TOC including EA	25th to 30th July 2020	5 days
	Evaluator	Evaluability Assessment – Plausibility, Feasibility, Credibility and Utility. Analysis of data and information and drawing conclusions		

Phases	Responsibility	Task	Proposed timeline	No.of man days
Finalisation of Report	Evaluator	Presentation of draft findings, preliminary lessons, best practices and recommendations. Finalisation of the Evaluability Assessment Report incorporating Recommendations	1st August 2020	3 days

4. Adherence to guideline and requirements—
   - Formulating and presenting recommendations
   - Identifying and presenting lessons learnt
   - Identifying and presenting emerging good practices

Evaluation manager should discuss the inception report with the evaluators until all points are mutually understood and agreed upon.

10.1.2 Final Report

The final report provides key issues, overall assessment, findings and conclusions, lessons learnt, emerging good practices, follow-up actions and recommendations.

1. Title Page
   - Title cited in the TOR
   - Type of evaluation (Independent/internal review)
   - Indication of time for evaluation (Mid-term, Ex-post or Impact)
   - List of geographical areas covered
   - Date of evaluation
   - Date of when the report officially ends
   - Name of the Evaluation Consultant
   - Name of the Evaluation Manager
   - Office Administrator for the evaluation
   - Donor and Project Budget
   - Cost of the evaluation

2. Table of contents
   - Contents of the report, List of Figures and Charts, Acronyms and Abbreviations
   - List of appendices includes TOR, Logical Framework, Inception Report, List of persons interviewed

3. Executive Summary
   - Executive summary (1-page high level summary and a 3-page executive summary)
   - Explanation of the project purpose, logic, objectives and background
   - Overview of the purpose, scope, clients of evaluation, time period, geographical coverage and beneficiaries
   - Concisely describe the evaluation methodology
   - Summary of evaluation findings
   - Concise list of conclusions, lessons learnt and emerging good practices and recommendations.

4. Body of the report
   - Project background
     - Brief outline of economic, social, political, cultural, historical context of the country.
     - Describe project objectives
     - Describe the context and intervention logic of the project
     - Project funding arrangement and donor
     - Organisational arrangement for project implementation
     - Role of donor, partners and other stakeholders
     - Project implementation—major events and milestones

   - Evaluation background
     - Purpose and primary use of the evaluation
     - Scope of evaluation (geographical coverage, phases, focus areas, gender etc)
     - Clients of evaluation and main audience of report
     - Evaluation criteria and questions that client want Evaluator to answer
     - Dates, events and operational sequences of the evaluation

   - Methodology
     Theory of change [Overview of the existing theory of change]
     - Evaluation Design Matrix
     - Evaluation Criteria and Key Evaluation Questions (KEQs) – OECD/DAC
     - Evaluation methods, data collection instruments and their justification
     - Description of the source of information and data used
     - Limitations and potential biases if any
     - Rationale for stakeholder participation in evaluation
     - Report includes adherence to evaluation norms, standards and ethical safeguards

   - Main findings
     Overall assessment of the project performance including relevance, coherence, efficiency, effectiveness, impact and sustainability.
     - Findings (supported by evidence)
     - Evaluation questions are adhered and explanation provided.
     - Disaggregated data by sex, age, ethnic group
     - Unintended or unexpected results
     - Factors contributing to successes and failures of programmes/projects
     - Crosscutting issues—Gender etc.

The findings should be supported with evidence and should include intended and unintended outcomes. Findings should be relevant to the purpose and scope of the project evaluation and all KEQs should be addressed.

5. Conclusions
   - Conclusions are formulated by synthesising the main findings into a statement of merit or worth
   - Conclusions
   - Judgements are fair, impartial and consistent with the findings
   - Validity and reality of conclusions

6. Lessons Learnt and Emerging Good Practices
   To improve project performance and organisational learning and general lessons learnt

   Lessons Learnt
   - Are the lessons significant? (Not trivial matters)
   - Lessons capturing the context
   - Use of lessons possible in the future

   Emerging Good Practices
   - Identifying successful practices
   - Describe how it works
   - Is it applicable in a different context?
   - Is it replicable in a different context?

7. Recommendations
   Recommendations should follow from conclusion and good practices
   - Specify who is called upon to act. Example: project management or donor?
   - Specify action needed to remedy the situation (Actionable recommendations)
   - The main findings should be set out priority-wise in the report as high, medium and low priority.
   - Specify the recommended time frame, follow-up and resource implications of the recommendations.

- Annexures
  - TOR
  - Inception Report
  - Evaluation Design Matrix
  - Logical Framework Analysis (LFA)
  - List of persons interviewed.

- Other additional matters to be considered in preparing the report
  - Developing actionable recommendations
    A key challenge in many countries is that the recommendations are not effectively put into use. It is important to ensure that the recommendations are evidence-based, actionable, practical and quality.

Some recommendations are not used due to lack of stakeholder involvement, low priority, not practical, not agreed by the implementing agencies (disagreement), not technically valid or factually incorrect, negative political implications, budget constraints and legislative implications.

It is important to ensure that the actionable recommendations are relevant to the KEQs and purpose of evaluation; based on evidence-based findings; technically feasible and implementable; relatively few in number; identify the responsible authority and be financially and operationally feasible to implement.

Key elements in the evaluation report: Utility perspective

The key elements of an evaluation report from a utility perspective cover: (i) overall assessment, (ii) key issues, (iii) lessons learnt and (iv) follow-up actions and recommendations

Overall Assessment: Overall assessment should provide a rating of the project. Donors such as ADB use a 4-band rating which is given below:

- Highly Successful (HS)
- Successful (S)
- Partially Successful (PS)
- Unsuccessful (US)

The rating is based on the DAC Evaluation Criteria and considered the evaluation questions with scoring and rating. The Evaluation criteria includes relevance, coherence, efficiency, effectiveness, impact and sustainability. Economic Internal Rate of Return (EIRR) should be at least 12% to be rated as highly successful.

Key Issues: Project related key issues in planning and Design, Implementation and post implementation (Operational Performance) should be briefly and clearly stated.

Lessons Learnt: Lessons learnt from the entire project cycle including planning and design, selection of consultants, participation of beneficiaries, delay in implementation and post implementation lessons for improvements of future projects should be clearly and briefly highlighted.

Follow-up actions and recommendations: This includes actionable recommendations required for the donor, borrower and executing agency on project-specific matters should be set out clearly. The most significant recommendations should be stated, rather than trivial matters and the responsibility for execution of the recommendation should also be stated.

Structure of the evaluation report:

In order to encourage the effective use of the evaluation report, it is structured into three sections to cater the different hierarchical levels of senior managers, operational staff and users. The report is structured as 1/3/25. This implies:

- 1-page policy level summaryfor Ministers/ Secretaries/very senior policy level officers
- 3 pages Executive Summary—for senior level Managers
- 25 Pages main Report—for operational level Managers and users

10.2 Dissemination and Feedback

The utility of any evaluation is a prime criterion for judging its worth, regardless of the technical, practical and ethical merit. It is important to make the evaluation report effective by ensuring its usage. For evaluation to have an impact, it needs to be disseminated and communicated to a wider audience, including primary and secondary users, to ensure behavioral changes and actions.

It is important to ensure that the evaluation report and its findings, lessons and recommendations reach multiple constituencies. These users can be broadly categorised as:

(a). Primary target group/Key actors: The primary target group includes those who are expected to take action on the findings (Eg. Planners, Policy Makers, Donors).

(b) Secondary target group/Influence Actors: The secondary target group includes those who influence the key actors (Example—Watchdog agencies such as media, CSOs, NGOs)

Characteristic of an Effective Feedback

- Responsive to the needs of different audience (Demand driven, client orientation, tailoring evaluation report to audience needs)
- Timely; Accessible; Simple (user friendly)
- Avoid Information Overload; On-going and systematic
- Use different approaches; Promote follow-up

Evaluation Feedback

This is the branch of evaluation which has so far received the least attention. Although a large stockpile of evaluation reports is prepared and available, they have not been read and acted upon. The main problem in the area of feedback relates to feedback at the policy level (Cracknell). For entities that approve projects, the Evaluation findings are seen as very useful, as it helps them to formulate successful projects based on the lessons and experiences of past projects. When donor-led evaluations are conducted, the feedback to the beneficiary country is neglected and it is not applied and used due to lack of ownership. Holding seminars, workshops, circulation of reports, synthesis reports and computerised evaluation information systems and evaluation briefs help to improve the use of evaluation findings.

Feedback also requires different kinds of skills, and communication skills are more especially needed than analytical skills. Eleanor Chelimsky states "The courage to say what users may not want to hear is the characteristic of an evaluation function that keep institutions honest". One is more likely to lose friends than to gain them.

Unfortunately, the evaluation reports often remain unread and no action is taken. It is important to engage the primary users of the report at the inception of the evaluation, so that their interests and needs could be addressed in the evaluation. Similarly, the reports have to be prepared in a user-friendly manner, to encourage usage. It is generally said by many practitioners that evaluation should not be done at all if there is no prospect for it being used and useful to the audience.

In order to generalise findings that recurr in several evaluations, they are captured through evaluation synthesis. Evaluation synthesis is more justifiable and policy-makers can be more easily convinced than with one-off project evaluations. Tailoring evaluation reports to the different needs of the users is also practiced as a way of effectively reaching an audience. Feedback also needs to have actionable recommendations and should reach the decision-makers at the time of taking the decision.

The 1/3/25 reporting structure and arrangements also help to reach different hierarchical levels of policy-makers and managers.

Mechanism to promote Evaluation Feedback

- Dissemination Mechanisms: The dissemination mechanism is the means by which evaluation information is prepared and directed to client groups, to ensure the effective use of the evaluation findings, lessons and recommendations.
- Institutional Mechanisms: These describe the way in which evaluation units are linked to other parts of the Ministries and agencies and how evaluation findings are formally considered and linked with such agencies. This formal institutional linkage and mechanism can contribute significantly to ensuring the effective feedback of evaluation findings.

10.2.1 Dissemination Mechanisms

There are various methods of dissemination mechanisms used by evaluation Commissioners to ensure that the findings, conclusions, learnings and recommendations are used and put into practice

- Abstracts and Summaries: A short description of the project or activity evaluated and a synopsis of the key findings and issues, conclusion, recommendations made and lessons learned.
- Annual Reports: Provide an opportunity to highlight findings, trends and lessons to synthesise recent experiences from a number of studies and to direct management attention to particular needs.
- Synthesis Report: Lessons Learned over a number of years in particular sectors (eg, health, roads) or on cross-cutting issues (eg. women in development, the environment) can be synthesised and the generalised findings and recommendations can be used as a powerful learning process.
- Bibliographies: This includes the listing of evaluation reports manually or in computerised databases, to enable users to review, read and act on findings which are of interest.
- Automated Systems: Automated databases which increase the accessibility of information from evaluations for various users, including project designers, policy and planning staff, management and evaluators themselves. An Evaluation Information System (EIS) not only captures project-wise evaluation lessons and findings but also provides sector synthesis to support policy-makers.
- Feedback Seminars: Feedback seminars promote the discussion of issues and lessons arising out of evaluations and help to understand what works, when and in what context. They can also strengthen knowledge management.
- Press Conferences and Presenting Evaluation Results: These are captured by the media and enable the sharing of lessons and emerging good practices and key issues widely, among many users.

10.2.2 Institutional Mechanisms

- Linkages to Senior Management: Evaluation reports are submitted to senior management and in some cases higher authorities for formal approval, which enables them to review the report and understand the findings and lessons.
- Linkages to Policy Development: The Linkages between evaluation agencies and policy units enables them to build relationships that help them to use the evaluation findings and lessons. Evaluation staff are invited by the policy unit for discussions to share evaluation experiences and evidences and use them in the policy formulation process.
- Linking to Planning: The establishment of a formal link between evaluation and project concept developing agencies such as the National Planning Department helps to use past project evaluation findings and lessons in formulating new projects and programmes.
- Linkages to Program Management: Project approval documents must indicate how evaluation lessons have been taken into account in the design, planning and management of interventions.
- Linkages to Agency Procedures: Lessons from evaluation experiences are incorporated into future agency policy and practices (eg. appraisal procedures) to improve the planning, design and appraisal procedures.
- Linkages to Training: Lessons from evaluations are extensively shared and used in training programs (e.g. training on project design) to improve knowledge to implement best practices and success stories.

10.3 Evaluation Synthesis

- Evaluation synthesis is a sector-based summary of the evaluation findings and lessons drawn from project evaluation experiences. A one-off single project evaluation is not strong enough to convince policy-makers and planners to generalise findings. Single evaluation findings may be challenging or unexpected, and are unlikely to be acted upon.

- Synthesis of many projects in a sector or cluster helps to generalise findings which provide strong and justified evidence. Hence, such cluster/sector wise findings are capable of influencing policy-makers, as the findings repeatedly appear in many single project evaluations.
- The synthesis enables the translation of technical findings into policy through effective communication. The synthesis enhances the ability of evaluation to feed into planning, budgeting and policy-making processes.

10.4 Management Response and Improvement Plan

The evaluation report that captures findings, key issues, lessons learnt and follow-up actions and recommendations should be technically valid, factually correct and approved by the evaluation steering committee. However, it is important for the executing agency, implementing agencies and the project office to agree with the findings, lessons learnt and recommendations. Therefore, an opportunity is provided to the executing agency, implementing agencies, project office and other relevant ministries, agencies and departments to provide a formal management response to indicate which recommendations they accept or disagree with. The disagreements should be justified, with reasons given. For this purpose, the commissioner of the evaluation should prepare a table listing out and summarising the recommendations prescribed in the evaluation report and ask the concerned department and ministries to indicate their agreement or disagreement on the recommendations. The management response is the document which comprises the recommendations of the evaluation report and the responses to those recommendations by the agency responsible for the implementation. The management response should be completed within a short time frame. The recommendations should be clear, concise and listed in a priority order when presented to the targeted audience—the office or unit responsible for its follow-up. The format used is set out below:

Figure 35: Management Response

Recommendations	Agreed/Disagreed	Reasons for disagreement

Improvement Plan

Once the recommendations have been agreed in the management response, the Departments are expected to draw-up an improvement plan on how the recommendations will be implemented. The purpose of the improvement plan is to ensure that the recommendations and findings of the evaluations are utilised by the relevant responsible agencies.

An improvement plan is prepared by the evaluation commissioners requesting the responsible department and agencies to submit a progress report on the implementation of the agreed recommendations set out in the management response.

The improvement plan will include the agreed follow-up steps and actions with timelines and responsibilities. The template for improvement plan is set out below.

Figure 36: Improvement Plan

Agreed Action	Priority Low/Medium/High	Responsibility	Deadline (Time)	Current Status of Progress

The Commissioner of the Evaluation and the steering committee is responsible for tracking the implementation of the actions set out in the improvement plan. This process enables the Commissioners and Steering committee to ensure that the evaluation findings, lessons and recommendations are effectively used.

10.5 Strategies to promote better use of evaluation

1. Utilisation-focused evaluations: The **Needs of the clientele,** especially the policy-makers, planners and decision makers should be given due consideration when designing the evaluation questions. This will ensure that the needs of the clients are addressed through the evaluation and as such, the evaluation findings and recommendations will be utilised effectively, with ownership.

2. Engagement of Stakeholders: Ensure the **buy-in** of all relevant stakeholders at the beginning and throughout the evaluation, through the **active engagement** of users and all stakeholders.

3. **Involve planning, budgeting,** programme management and policy-making functions in designing, implementation and post implementation of the evaluations, to ensure the effective use of the evaluation findings.

4. Provide **credible and independent evidence to improve the quality and trustworthiness of the evaluation findings, which will enable a wider use of the evaluation findings.**

5. Emphasis on the **Management Response** and the **Improvement Plan** will ensure that the accepted recommendations are implemented in time, by the responsible agencies.

6. **Disseminating the findings** through workshops, seminars, synthesis reports and Evaluation Information systems (EIS) helps to widely share the recommendations to enable the use of evaluation lessons and findings.

7. Incorporating the evaluation lessons into Project concept documents and **Project Submission** Forms helps to formally link evaluations with the formulation of new projects and programmes.

8. The shift in focus from **donor-driven evaluations** to **country-led evaluations** and donor-partner joint and collaborative evaluations will enhance ownership of the evaluation findings, which in turn facilitates the effective use of the findings and recommendations.

9. Evaluation reports when tabled to Cabinet of Ministers**,** and Parliament and made available to the public, **it helps** to use the evaluation findings and recommendations.

10. When placed in the **National Evaluation Information System (EIS) (Repository) evaluations helps to improve knowledge management and can facilitate the effective use of evaluation findings.**

11. The National Evaluation Policy makes evaluation mandatory and ensures its use.

These are some of the strategies to promote the effective use of evaluation findings and recommendations.

References

1. **V. SIVAGNANASOTHY, (2013) Country-led National Evaluation System: Independence, Credibility and Use of Evaluations; Challenges and Solutions, 3rd UN International Conference , National Evaluation Capacity ,Conference Publication**

2. **V. SIVAGNANASOTHY, (2011) - Use of Evaluation in Public Policy ,Sri Lankan Experience , 2nd UN International Conference, National Evaluation Capacity, Conference Publication**

3. V. SIVAGNANASOTHY. (2010) The National Monitoring and Evaluation System in Sri Lanka Experiences, good practices, challenges and the way forward, from Policies to Results , Unicef Publication

4. Velayuthan Sivagnanasothy (2016) (Main Author), Sri Lankan Paper on Collaborative Partner-Donor Evaluation ;This study was fed into the Study on Collaborative Partner –Donor Evaluation –Final Study Report – OECD Publication

5. Velayuthan Sivagnanasothy (2009), Sri Lanka : National Monitoring and Evaluation System : Experiences, Challenges and the Way Forward – 1st UN International conference on NEC, UNDP Conference Publication

6. OECD / DAC Evaluating Development Cooperation : summary of Key Norms and Standards (www.oecd.org/dac/evaluationnetwork)

7. UNEG (2016), Evaluation Competency Framework (www.unevaluation.org/2016-Evaluation-Competency-Framework.)

8. Independent Evaluation Group, Asian Development Bank (2006), Guidelines for Preparing Performance Evaluation Reports for Public Sector

9. The Presidency, Republic of South Africa (2011), National evaluation Policy Framework (www.thepresidency.dpme.gov.za)

10. IEG ,World Bank , Impact Evaluation – The Experience of the Independent Evaluation Group of the World Bank,World Bank Publication.

11. Basil Edward Cracknell (2000) Evaluating Development Aid – Issues, Problems and Solutions, SAGE Publication

12. Jody L Fitzpatrick, James R Sanders, Blaine R Worthen (2011), Programme Evaluation – Alternative Approaches and Practical Guidelines , Pearson Publication (4th Edition)

Notes

Notes

Notes

This publication is made possible by the generous support of the American People through the United States Agency for International Development (USAID) funded Strengthening Democratic Governance and Accountability Project (SDGAP). The content of this manual is the sole responsibility of the authors and does not necessarily reflect the view of USAID or the United States Government.

This publication has been developed with assistance from

USAID
FROM THE AMERICAN PEOPLE

Strengthening Democratic Governance
and Accountability Project (SDGAP)
in Sri Lanka