



# Guidelines on Evaluation in the National Government



# Version Notes

Released in 2020, the maiden version of these Guidelines has been developed in response to the directives under the National Evaluation Policy Framework (NEPF) and at the request of national government agencies (NGAs) for more guidance on the substance and processes of evaluation.

This “Version 1.0” of the NEPF Guidelines is preliminary and for demonstration and testing by a select set of national government agencies (NGAs). NEDA and DBM thought it would be more productive to first subject these Guidelines to demonstration before government-wide effectivity. After all, these Guidelines are intended to be a “living document” that will be updated regularly based on feedback from stakeholders and housed in an interactive National Evaluation Portal (currently under development).

Version 1.0 leaves out sections originally included in drafts that were circulated to government agencies in 2018. These sections require substantive inputs and have been set aside temporarily for further study. In particular, there has been a clamor for guidance on institutional arrangements for evaluation. Rather than impose boilerplates for the institutional arrangements, organizational structure, and staffing pattern for all agencies, an accounting of the unique needs and contexts of agencies is necessary to better provide appropriate guidance and support.

## Copyright

Published July 2020

All rights reserved. Any part of this publication may be used and reproduced, provided proper acknowledgement is made.

Produced by the National Economic Development Authority, and the Department of Budget and Management, with technical support from United Nations Development Programme (UNDP) in the Philippines, for the Strategic M&E Project.

*UNDP is the leading United Nations organization fighting to end the injustice of poverty, inequality, and climate change. Working with our broad network of experts and partners in 170 countries, we help nations to build integrated, lasting solutions for people and planet.*

*Learn more at [undp.org](https://undp.org) or follow at [@UNDP](https://twitter.com/UNDP).*



# Foreword

Recognizing the importance of results-based monitoring and evaluation in improving planning, programming, budgeting, and implementation of the government's development programs and projects, the National Economic and Development Authority (NEDA) and the Department of Budget and Management (DBM) issued Joint Memorandum Circular No. 2015-01 in July 2015, which established the National Evaluation Policy Framework (NEPF).

The NEPF aims to embed an evaluative culture in the government through actual practice of evaluation in support of good governance, transparency, accountability, and evidence-based decision-making in the public sector. To this end, the NEPF sets the standards and guidelines of evaluation in the government. To build evaluative institutions, the NEPF promotes monitoring and gathering evidence to gauge progress with respect to the intended results and outcomes of the Philippine Development Plan (PDP).

As NEDA oversees the implementation of the PDP, it also leads the systematic, rigorous, and impartial assessment of its progress. This entails data gathering, analysis and reporting on the effectiveness of the policies, projects, and programs in achieving desired results. Monitoring and evaluation (M&E)

reports are particularly useful to DBM, in budget allocation and public resource management. M&E complements reforms in policies, information systems development, and engagement with civil society organizations, which are important for improving delivery of public goods and services.

Following the issuance of the NEPF in 2015, pilot evaluations and capacity building have been conducted. The NEDA and DBM collaborated on a set of guidelines that will standardize the content and quality of evaluation products, outline and streamline the process for each stage of evaluation, and define the roles and responsibilities of key institutional and individual stakeholders. The set of guidelines, which is now ready for roll out, is also a response to the clamor for guidance on the preparation, design, conduct, management, and utilization of evaluation studies in the country. As such, it shows the ways to prepare an evaluation plan and evaluability assessment and how to carry out quality assurance, facilitate management responses, and disseminate and utilize quality results.

To adapt to the needs of the times, this initial version of the NEPF Guidelines will need subsequent enhancements through feedback from stakeholders. Thus, we enjoin the support of national government agencies, state universities and colleges, government-owned and/or controlled corporations, and government financial institutions, as well as public-private partnerships, in integrating the practice of evaluation within their respective mandates. We look forward to your cooperation and participation in enhancing this document as we embed on evaluative thinking and improvement in the way we formulate, design, and implement our policies, programs, and projects – ultimately to make development work for the people.

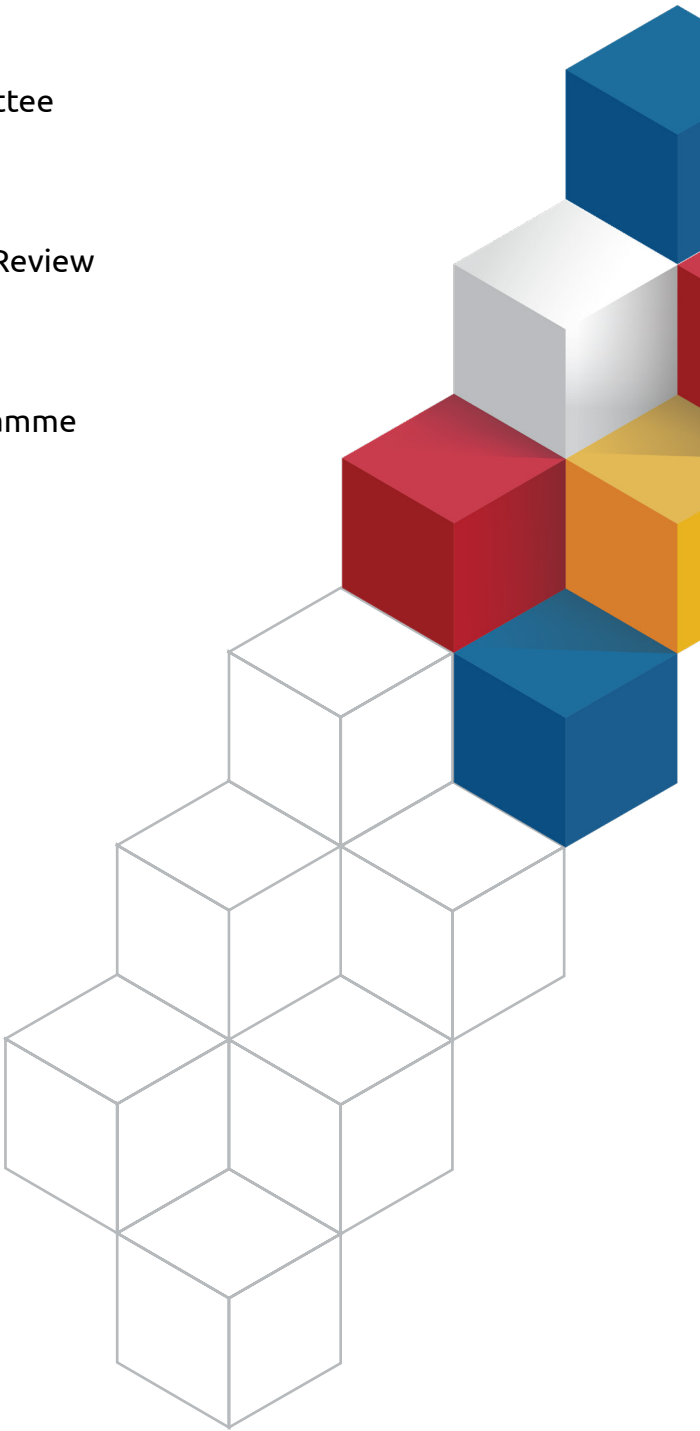
**Karl Kendrick Chua**  
Acting Secretary, National Economic and  
Development Authority

**Wendel Avisado**  
Secretary, Department of Budget and  
Management

# Abbreviations

AO	Administrative Order
BPF	Budget Priorities Framework
DBCC	Development Budget Coordination Committee
DILG	Department of the Interior and Local Government
DBM	Department of Budget and Management
ETF	Evaluation Task Force
ERG	Evaluation Reference Group
FOI	Freedom of Information
GOCC	Government-Owned and -Controlled Corporations
GFI	Government Financial Institution
HEI	Higher Education Institution
M&E	Monitoring and Evaluation
NGA	National Government Agencies
NEDA	National Economic and Development Authority
NEPF	National Evaluation Policy Framework
ODA	Official Development Assistance
OECD-DAC	Organisation for Economic Co-operation and Development-Development Assistance Committee
OP-PMS	Office of the President - Presidential Management Staff
OPIF	Organizational Performance Indicator Framework
PDP	Philippine Development Plan
PDP-RM	Philippine Development Plan-Results Matrix
PIDS	Philippine Institute for Development Studies
PIP	Public Investment Program
PREXC	Program Expenditure Classification
PSA	Philippine Statistics Authority
PSM	Public Sector Management

RA	Republic Act
RBMER	Results-Based Monitoring, Evaluation, and Reporting
RBPMS	Results-Based Performance Management System
RDC	Regional Development Council
RDP	Regional Development Plan
RPMC	Regional Project Monitoring Committee
SDGs	Sustainable Development Goals
SER	Socioeconomic Report
SEER	Sector Efficiency and Effectiveness Review
SUC	State Universities and Colleges
ToR	Terms of Reference
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
WB	World Bank



# Contents

<b>I. INTRODUCTION</b>	10	<b>IV. PREPARING AN EVALUATION</b>	46
1.1. What is an evaluation?	13	4.1. Program objectives and intended outcomes	46
1.2. Why evaluate?	13	4.2. Evaluation objectives	48
1.3. Evaluation in the context of government’s drive towards results	16	4.3. Evaluation methods	53
1.4. Evaluation, development planning, and investment programming	18	4.4. Risk identification and assessment	58
1.5. Evaluation and the budget cycle	22	4.5. Financial and human resource requirements	61
1.6. Evaluation principles, norms, and standards	24	<b>V. IMPLEMENTING EVALUATIONS</b>	64
1.7. Applicability of guidelines	29	5.1. Evaluation governance arrangements	64
1.8. Overview of the evaluation process	30	5.2. Managing risks	69
<b>II. INSTITUTIONAL RESPONSIBILITIES</b>	33	5.3. Ethical behavior	70
2.1. Implementing agencies	33	5.4. Evaluation reporting	71
2.2. Evaluation Task Force and Secretariat	35	<b>VI. UTILIZING EVALUATIONS</b>	76
2.3 Role of Regional Offices	37	6.1. Management response	77
<b>III. INITIATING AN EVALUATION</b>	38	6.2. Communication and dissemination	78
3.1. Developing an agency evaluation agenda	38	6.3. Planning for the next evaluation	80
3.2. Developing an evaluation plan	39	<b>VII. QUALITY ASSURANCE</b>	81
3.3. Assessing evaluability	40	<b>REFERENCES</b>	83
3.4. Evaluations vs. internal assessments	43	<b>ANNEXES</b>	87
3.5. Stakeholder engagement	44		



# Introduction

- A. Countries across the world agree that monitoring and evaluation play a strategic role in helping governments make informed decisions that can transform public services, and, consequently, people's lives. While a few countries have made significant headway in legislating and setting up their respective whole-of-government systems for evaluating major programs and projects, most other countries have either set up fragments of evaluation mechanisms for certain sectors like health and education; or been conducting evaluations though merely for specific programs, in compliance with donors' requirements for sustained funding (Rosenstein, 2015).
- B. The adoption of the United Nations General Assembly Resolution A/RES/69/237 in 2014, and of the Sustainable Development Goals (SDGs) in the following year, created a solid consensus to strengthen national capacity for evidence-based decision-making, while opening opportunities for most countries to level off with frontrunners. Particularly, the Bangkok Declaration in October 2015 sealed the commitment of more than 100 countries to take collective action to support their respective evaluation capabilities in light of meeting their country-level SDG targets by 2030 (UNDP IEO, 2015).

C. Furthermore, the Global Evaluation Agenda 2016-2020 (EvalPartners, 2016) has set to motion a framework that spells out strategies for countries to reach their full evaluation potential in response to the SDG's "no one left behind" principle:

- Create an enabling environment through policy, resource allocation, and systems to ensure access to results;
- Build institutional capacity of public sector agencies and their partner institutions;
- Strengthen individual capacity of evaluation commissioners, users, and implementers to meet the demand for these roles; and
- Support inter-linkages among the first three strategies.

D. In July 2015, the National Economic and Development Authority (NEDA) and the Department of Budget and Management (DBM) released the Joint Memorandum Circular No. 2015-01, which defines a national evaluation policy framework (NEPF) applicable to all government entities and instrumentalities across the country. The objectives of the NEPF are to:

- Support evidence-based decision-making for programming by providing evidence on the outcomes attributable to programs, efficiency with which outcomes are achieved, and extent to which outcomes align with national priorities
- Enable continuous program improvement by providing feedback and learning to managers and other stakeholders on program performance
- Ensure public accountability and transparency through the public disclosure of all evaluations

E. The NEPF is designed to deepen the implementation of ongoing efforts to improve public sector performance through evidence, including: (i) the use of results matrices to monitor progress towards achieving the outcomes of the Philippine Development Plan (PDP) 2017-2022; (ii) outcome-based, performance-informed budgeting initiatives of DBM; and, (iii) a Results-Based Performance Management System introduced in 2011 to assess and measure the performance of government agencies. Recent reforms such as (i) the Program Expenditure Classification (PREXC) structure of the National Budget and (ii) the introduction of cash-based appropriations scheme are complemented by the implementation of the NEPF.

- F. In terms of scope, all programs and projects implemented by all National Government Agencies (NGAs), State Universities and Colleges (SUCs), Government-Owned and/or Controlled Corporations (GOCCs), and Government Financial Institutions (GFIs) with budgetary support from the National Government through local and foreign funds, as well as public-private partnerships, are to be reviewed at least once at the end of their program/project life cycle or as frequent as necessary. To optimize resources, programs and projects that meet the prioritization criteria as defined in these Guidelines will be subjected to an evaluation.
- G. The Guidelines on Evaluation in the National Government, to be subsequently referred to as the Guidelines, have been developed to translate the NEPF into action. The Guidelines provide a concise statement of the institutional responsibilities, rationale, processes and methods for implementing evaluation across the entire public service. The Guidelines should be read in conjunction with other manuals and guidelines issued on monitoring and evaluation<sup>1</sup>.
- H. The Guidelines have been prepared for: government staff who are involved in planning, commissioning, and managing evaluations; managers from implementing agencies who use evaluations to assure the relevance, coherence, efficiency, effectiveness, and sustainability of programs and projects; users of evaluation results: stakeholders and partners, the legislature, and oversight agencies; evaluators who need to understand government evaluation principles, standards, and processes; and the broader national development and evaluation community
- I. The Guidelines are organized into the following sections:
- Overview of evaluation in the context of results-based management
  - Institutional roles and responsibilities in evaluation
  - Initiating an evaluation
  - Preparing an evaluation
  - Implementing an evaluation
  - Utilizing the evaluation findings and recommendations
  - Quality assurance of an evaluation

<sup>1</sup> These include: NEDA. n.d. NEDA Ex-post evaluation manual. Manila: National Economic and Development Authority and NEDA. 2017 Results-based monitoring and evaluation (RbME) guidebook. Draft. Manila: National Economic and Development Authority.

## 1.1 What is an evaluation?



**Evaluation is the systematic, rigorous, and impartial assessment** of a program or project to ascertain its level of achievement. It examines the design, implementation, context, and results of an intervention based on its relevance, coherence, efficiency, effectiveness, impact, sustainability and other criteria.

To provide credible and useful information that organizations and stakeholders can use in their decision-making, evaluations must be conducted independently (OECD-DAC, 2002; UNEG, 2018).



**Evaluation aims to promote accountability and learning.** It seeks to understand why, and to what extent, intended and unintended results are achieved, and to analyze the implications of these results to programs and projects. Evaluation can inform planning, programming, budgeting, implementation, and reporting; and can contribute to evidence-based policymaking, development effectiveness, and organizational effectiveness.



Like the monitoring activities and internal assessments or reviews done by program implementers, **evaluation provides evidence to support the improvement** of program and project design and performance.

However, evaluations are much more rigorous and systematic in design and methodology; and hence must be conducted in the most impartial and objective manner possible by an independent evaluator. Independent evaluators may include consultants or consulting firms, academic institutions, and research organizations, among other third parties.

In contrast, internal assessments or reviews conducted by program or project staff, while providing valuable information, are not considered to meet the standards for impartiality that are prescribed in the NEPF and





these Guidelines.

An agency staff may be able to conduct an evaluation if the highest possible degree of impartiality can be ensured to **maximize the objectivity and credibility of the results**, and to minimize bias and avoid conflict of interest.

Ideally, an evaluator that is internal to the organization may be an individual or unit reporting directly to the head of agency or a governing board; or, at least, not belonging to the program or project design and implementation unit, and/or not reporting to an agency official who is responsible for leading the implementation of the program or project being evaluated.

All evaluations must be conducted in an independent and transparent manner. Reports, findings, and management responses must be made publicly available.





## 1.2 Why evaluate?

Evaluation plays an important role in the design and implementation of government programs, and in bringing effective development as a whole. It provides a systematic process for collecting credible data and using it to assess whether a program remains to be the best policy response for a specific issue or intended outcome (Queensland Treasury and Trade, 2014).

Evaluation can be used to assist agency program managers in understanding how programs should ideally work in order to achieve their targets, and hence guide them on assessing whether the program is doing the right things at the right time to achieve the right results (Queensland Treasury and Trade, 2014). On a larger scale, it can also help the government in assessing a program’s appropriateness and value for money, hence helping them in improving resource allocation (Ibid.).

If properly planned and implemented, evaluation helps in building an evidence base that ensures every peso spent yields results that can be seen, felt, and measured. Government agencies, public servants, and the community can all benefit from evaluation (see Table 1), particularly in how its findings can inform future development of programs and projects, and optimize the delivery of ongoing programs.

Table 1: Benefits of evaluation for stakeholders

Stakeholder	Potential benefits
 Government	<ul style="list-style-type: none"><li>■ Information to assist decision-making</li><li>■ Improved ability to achieve government priorities</li><li>■ Efficient resource allocation</li><li>■ Highlights achievements and opportunities to strengthen performance</li><li>■ Encourages greater public trust in government</li></ul>
 Agencies	<ul style="list-style-type: none"><li>■ Stronger basis for informing government priorities and resource allocation</li><li>■ Improved service delivery and client satisfaction</li><li>■ Builds an agency’s reputation for innovation and continuous improvement</li></ul>
 Public Servants	<ul style="list-style-type: none"><li>■ Develops new skills and broadens experience</li><li>■ More opportunity to shape public policy</li><li>■ Fosters a more dynamic and creative work environment</li><li>■ Recognizes and rewards efforts to improve performance</li></ul>
 Community	<ul style="list-style-type: none"><li>■ Better government services</li><li>■ Informative government reporting</li><li>■ Transparent and accountable government</li><li>■ Public monies used more efficiently</li><li>■ Greater confidence in activities of government</li></ul>

Source: Western Australia Government (2015)



# 1.3 Evaluation in the context of government’s efforts to strengthen results-based management (RBM)

A. The turn of the 21st century witnessed the introduction of reforms to intensify the Philippine government’s drive towards measurable results. In the context of public expenditure management reforms, NEDA and DBM jointly undertook the Sector Efficiency and Effectiveness Review (SEER) in 2001 to assess the responsiveness of programs and projects to sectoral objectives under the Philippine Development Plan (PDP).

In 2007, DBM introduced the Organizational Performance Indicator Framework (OPIF) to align the budget with the delivery of goods and services to achieving desired outcomes.

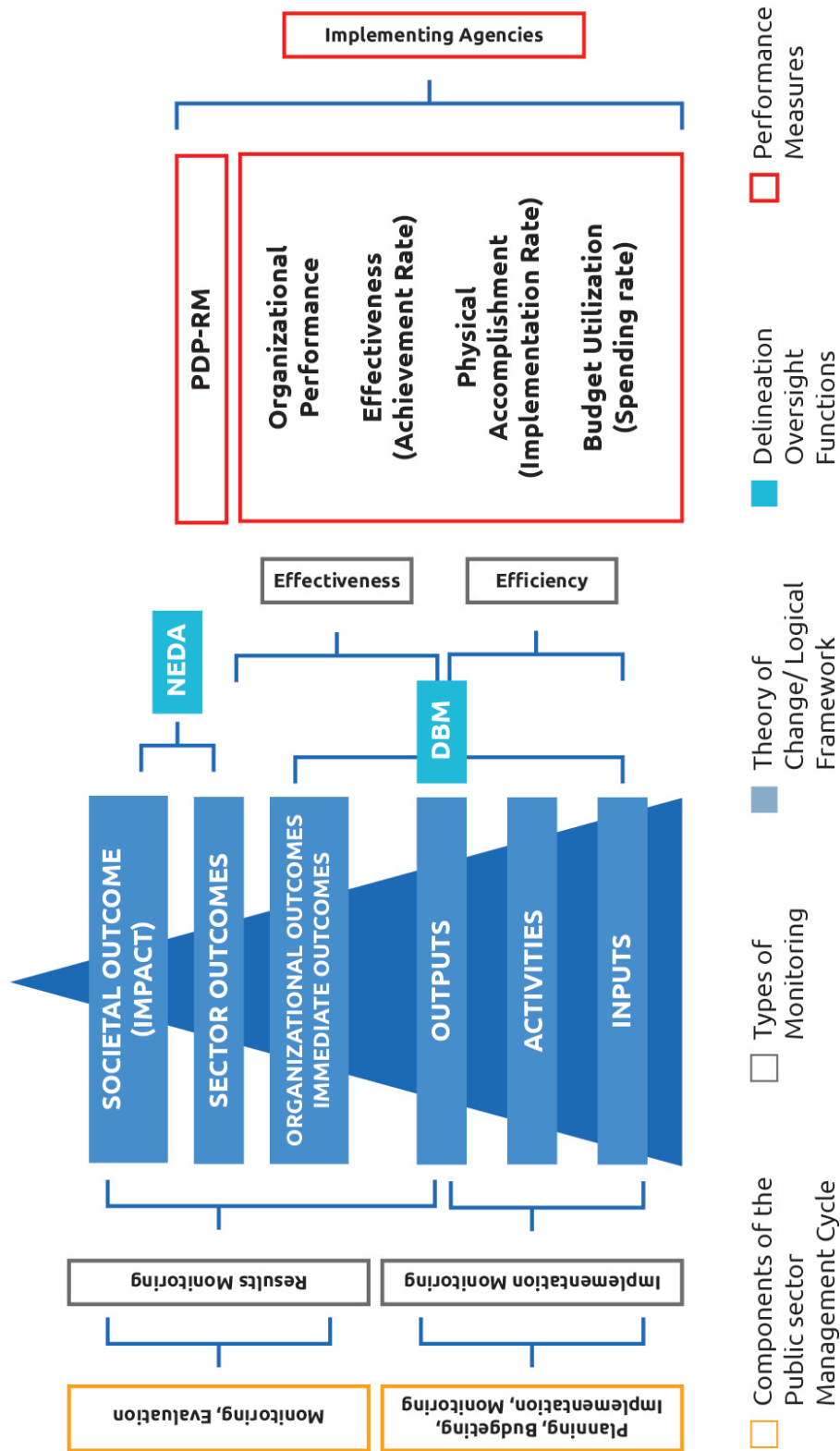
B. Under the auspices of Administrative Order (A.O.) No. 24 issued in 2011, the government introduced the integrated Results-Based Performance Management System (RBPMS; see Figure 1) to align the achievement of broader policy and sectoral objectives with organizational and individual performance.

The same year, NEDA introduced the PDP Results Matrices (PDP-RM) which spells out the outcome indicators and targets to be achieved during the PDP period. Building on OPIF, DBM in 2013 introduced the Performance Informed Budget (PIB) structure to build the output-level indicators and targets—and beginning 2014, indicators at the outcome level—into the Budget itself.

C. Since NEDA and DBM issued the NEPF in 2015, the government expanded efforts to reinforce a results orientation in the public sector.

In 2016, DBM introduced the Results-Based Monitoring, Evaluation, and Reporting (RBMER) Policy Framework “to strengthen, streamline, and standardize the monitoring, evaluation, and reporting [of performance information] to support policy and program improvement, expenditure management, and local and national decision-making.” (DBM,

Figure 1. RBM in the Philippine Government



Source: RBMER Policy Framework, DBM (2018).

2016). Building on OPIF and PIB, DBM rolled out the Program Expenditure Classification (PREXC) structure to organize activities and projects under programs that are assigned with both output- and outcome-level indicators and targets.

D. All these reforms strengthen the government’s results orientation and clarify responsibilities and accountabilities. These reforms provide a strong basis and need for the purposive conduct of evaluations on government programs and projects, especially those which are critical to the attainment of goals under the PDP and Sustainable Development Goals (SDGs). In the context of these reforms, evaluation can be maximized to ascertain if desired outcomes are truly being met through government interventions; as well as to identify and address issues in design and implementation to optimize the impact of these interventions.

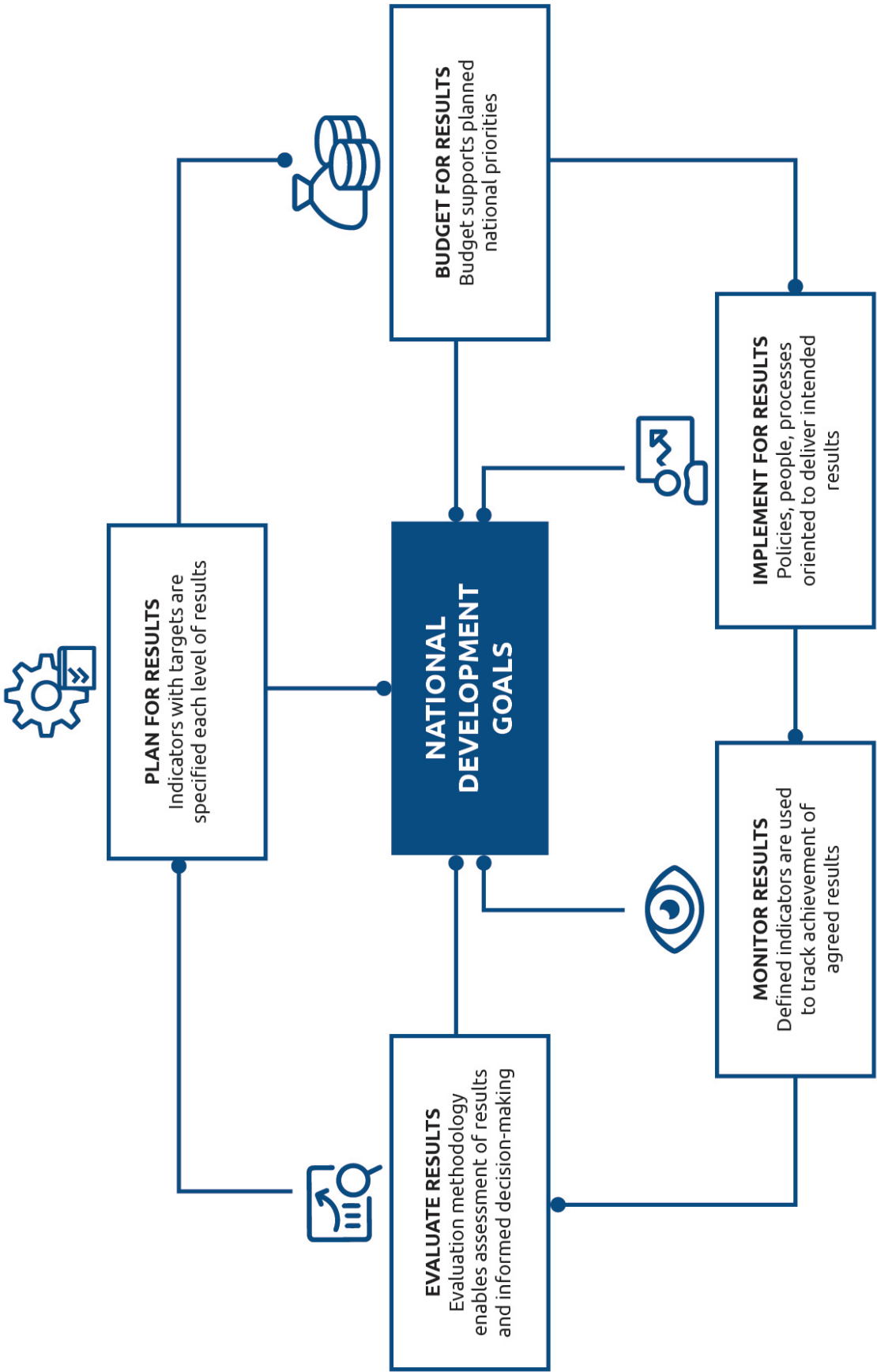
# 1.4 Evaluation, development planning, and investment programming

A. As an integral phase in the government’s Public Sector Management (PSM) cycle (see Figure 2), evaluations shall be guided by the priorities spelled out in, and shall inform the crafting and updating of, medium-term development plans and investment programs.

Evaluations shall be used to validate the results reported by agencies against their targets in the Philippine Development Plan (PDP) and its accompanying Results Matrix (RM). Evaluations shall also help in drawing out lessons and recommendations that could improve the implementation and maximize the impact of programs and projects under the Public Investment Program (PIP).

B. The NEPF mandates all implementing agencies to formulate their respective evaluation agendas, which shall align with the six-year period of the PDP, and, ideally, shall be updated annually to reflect emerging budget and investment priorities (see Section 3.1 for guideposts in developing an evaluation agenda).

Figure 2: Evaluation and the Public Sector Management Cycle



Source: APCoP (2011) as cited in NEDA (2018)

The Evaluation Task Force (ETF) created under the NEPF shall also develop a National Evaluation Agenda which spells out the evaluations to be commissioned by it. Compared to agencies’ evaluation agendas, which are typically focused on agency-specific programs or project, evaluations to be commissioned by the ETF shall focus on themes, sectors, or programs that concern multiple agencies; meta-evaluations of evaluations conducted by implementing agencies and other entities; and on specific priority programs that are critical to the success of the PDP and PIP (e.g., in terms of value of investment).

C. The results of evaluations shall feed into the annual Socio-Economic Reports (SERs) and Official Development Assistance (ODA) Reviews. Evaluation results shall likewise be used as inputs to the midterm updating of the PDP and PIP, and in the formulation of the succeeding PDPs and PIPs.

As the PDP and PIP are shaped by Ambisyon Natin 2040 – the country’s long-term vision for development – and the Sustainable Development Goals (SDGs), evaluations shall endeavor to inform the achievement of these long-term development objectives.

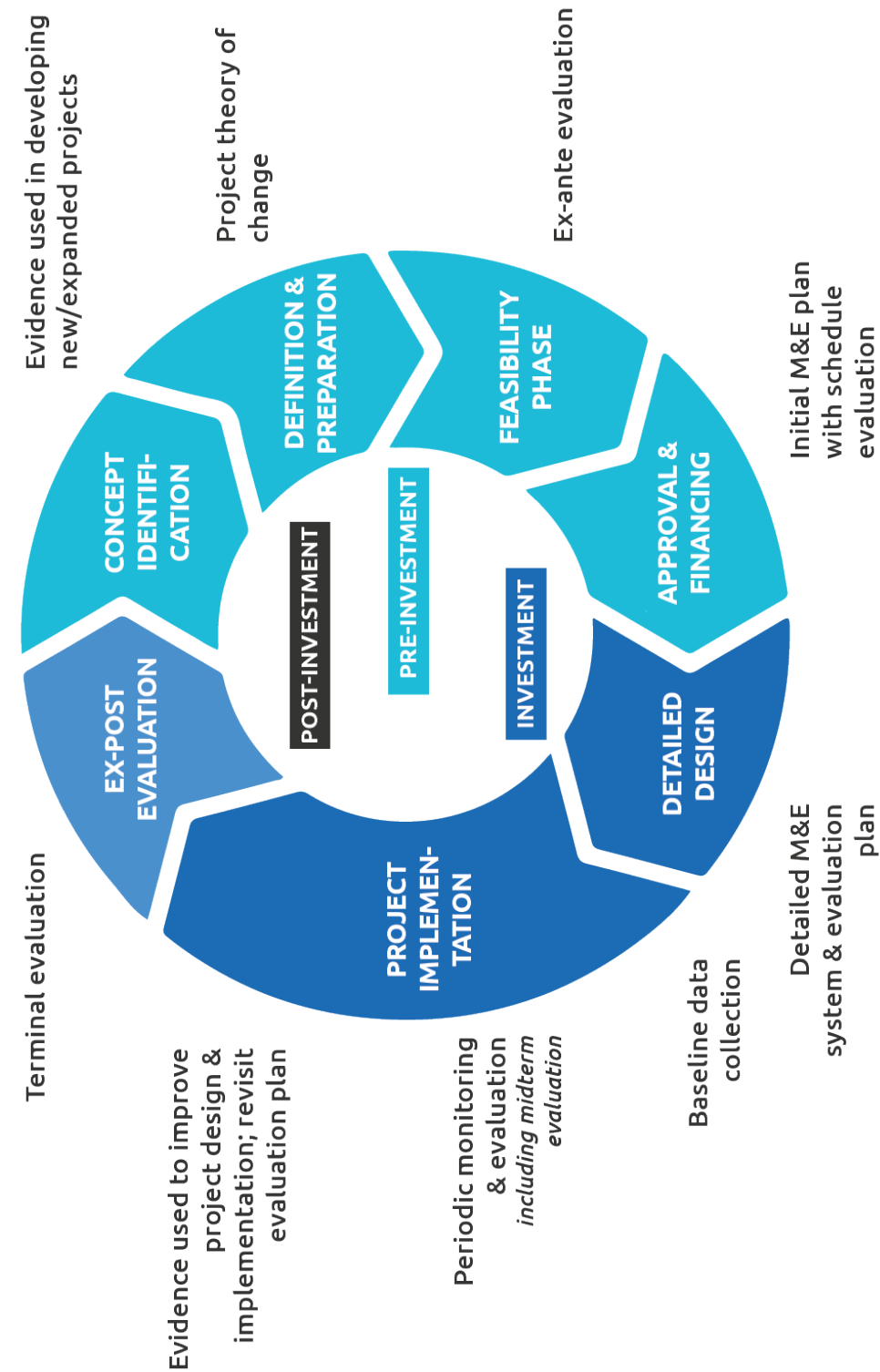
D. An evaluation will be most effective when spelled out and embedded as a part of a program or project at its onset, rather than treated as a separate process. Making evaluation thinking a function of program design and development results in a more proactive and regular use of evaluation findings, and, consequently, improved program delivery (Queensland Treasury and Trade, 2014).

The diagram below contextualizes evaluation within the project development cycle

E. To ensure evaluations are relevant, timely, costed, and sufficient in data, planning for an evaluation should start during program design. Likewise, where possible, baseline data should be collected prior to program implementation.

Evaluators and program implementers shall regularly provide feedback to ensure evidence is collected as needed, and the findings are useful enough to improve program delivery (Queensland Treasury and Trade, 2014).

Figure 3: How evaluation figures in project development and implementation



Source: NEDA (2005 & 2014); Queensland Treasury and Trade (2014). Notes: i) this illustration assumes the ideal situation that an M&E system and evaluation plan is crafted with the project. In case an ongoing project has no existing M&E system and/or evaluation plan, an evaluability assessment is necessary to determine the appropriate evaluation design given the availability of data and other requisites; ii) this project cycle is applicable to all national government projects, not just those appraised by the NEDA Investment Coordination Committee

# 1.5 Evaluation and the budget cycle

A. The National Budget translates the country’s medium-term social and economic goals into specific programs and projects to be funded annually, and which should contribute to achieving specific results at the output and outcome levels.

An evaluation shall enhance the preparation, legislation, implementation, and accountability of the Budget, and help tighten the linkage between planning, budgeting, and results (see Figure 4).

B. The NEPF likewise mandates agencies to create an evaluation plan for every program and put forward a project proposal for annual budgeting (see Section 3.2, Developing an evaluation plan). The program or project proponent shall likewise make use of prior evaluation of similar initiatives in designing the proposed program or project. Specifically, implementing agencies shall:






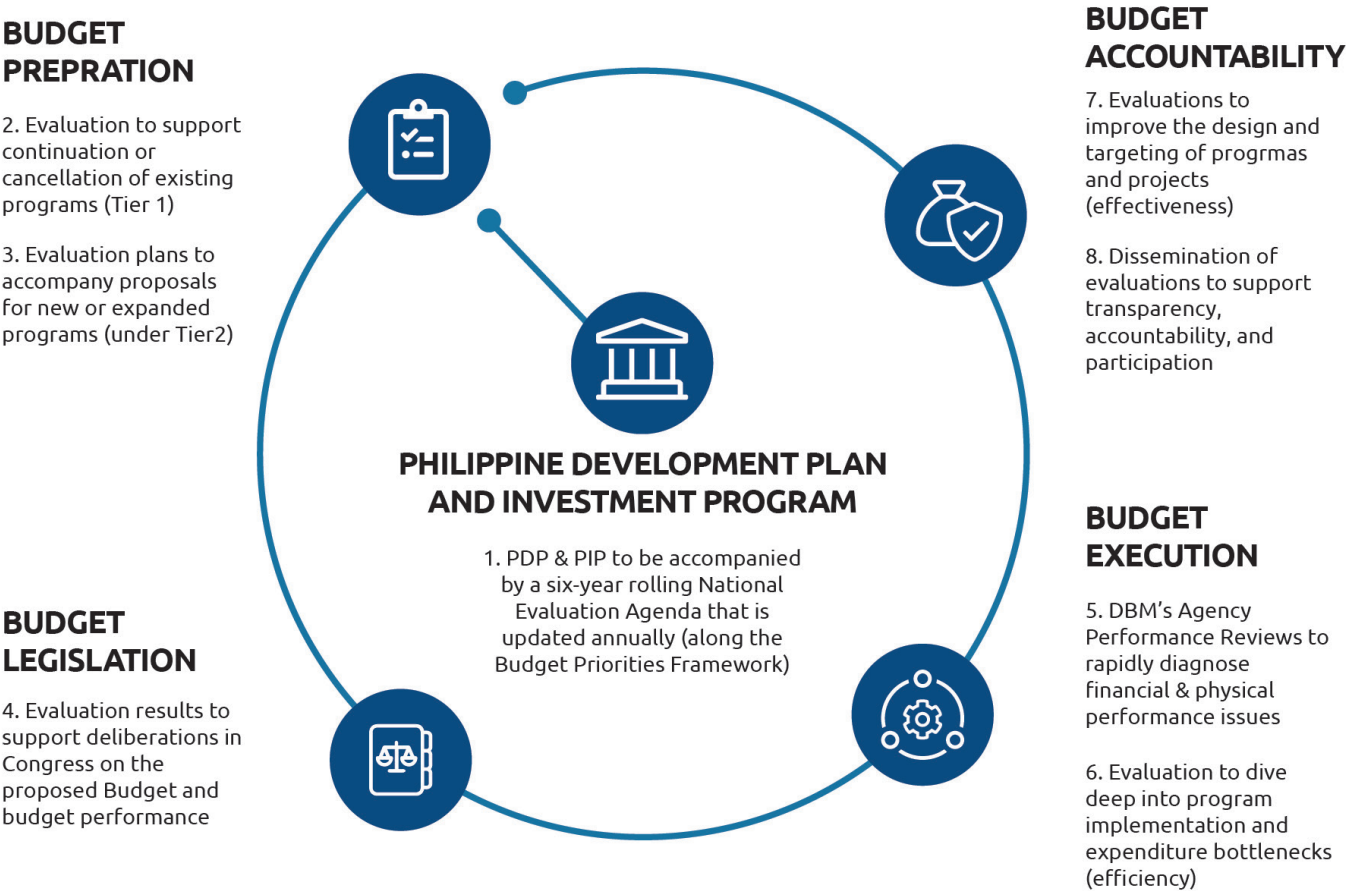
-  Ensure that new program or project proposals come with a costed evaluation plan, including the collection of baseline data
-  Use evaluation results to support decisions on whether to continue, expand, or terminate programs and projects, or to improve their design or implementation
-  Outline the proposed approach for evaluating any initiatives submitted for consideration as part of the budget process, using evaluation approaches that are appropriately designed to reflect the size, risk, and priority of the proposed initiative
-  Identify appropriate resourcing requirements, including the estimated cost, of the evaluation where it is deemed necessary or is given a high priority by the government
-  Incorporate the evaluation of newly-funded initiatives into an agency's operational plan.

Figure 4: Evaluation and the budget cycle

Source: DBM (2018)



C. To inform the formulation of the upcoming fiscal year’s Budget, evaluations should be completed by the first quarter of the preceding year in time for the deliberation of budget proposals for new or expanding programs (Tier II).



# 1.6 Evaluation principles, norms, and standards

Evaluations should uphold, promote, and contribute to achieving the goals and targets spelled out under Ambisyon 2040, the 2030 Agenda for Sustainable Development (SDGs), and the Philippine Development Plan (PDP). In relation, they should be guided by a set of generally accepted norms, standards, and principles as expressed and adopted in the local context through existing policies (UNEG, 2016; NEDA & DBM, 2015).



**Utility:** All evaluations should intend to come up with results that are relevant enough to influence decision making and improve service delivery; and accessible enough to facilitate learning and empower each stakeholder.



**Credibility:** An evaluation is credible when mechanisms and safeguards are in place to ensure rigorous methodology, inclusive process, and transparent reporting of results.



**Independence:** An evaluation’s independence further lends credibility to both its methodology and results. Independence comprises of two aspects:

- Behavioral, which entails protection from interference and the absence of undue influence by any party throughout the evaluation
- Organizational, which involves putting systems in place to ensure evaluation heads and staff are placed independently from policy, operational, and management units; are given full discretion on designing and implementing the evaluation; and are provided with sufficient resources to perform their roles (UNEG, 2006; WB, 2007). The independent conduct of evaluations creates an environment of impartiality, where potential conflicts of interest and opportunities for bias, political interference, or manipulation of results are identified and prevented.



**Impartiality:** Annex E of the NEPF underscores the importance of keeping evaluations fair—particularly, ensuring a balanced and comprehensive presentation of the program or project’s strengths and weaknesses. It places the responsibility to the evaluation units of implementing agencies to:

“ [E]nsure evaluations are conducted with the highest possible degree of impartiality in order to maximize objectivity and minimize the potential for bias. As a rule, it is appropriate to commission an independent third- party evaluator to undertake portions of, or the entire, evaluation.

While their active involvement is typically necessary to effectively conduct an evaluation, managers and other stakeholders of the program / project being evaluated should not be allowed to influence evaluation findings.

Good practices involve a constructive mix of internal and external stakeholders so that program knowledge can be brought to bear while potential conflicts of interest can be identified and prevented (NEDA & DBM, 2016). ”



**Ethics:** Evaluations shall adhere to standards of integrity and ethics. As fleshed out in Annex C of the NEPF:

- All those engaged in designing, conducting, and managing evaluation activities shall abide by the Code of Conduct and Ethical Standards for Public Officials and Employees (Republic Act No. 6713).
- Evaluators shall respect the right of the implementing agencies and individuals to provide information in confidence and ensure that sensitive data cannot be traced to its source.

Evaluators must ensure that those involved in evaluations have the opportunity to review and approve the statements attributed to them. In addition:

- Rights and safety of respondents and other subjects of evaluations, which includes ensuring the privacy and confidentiality of their data as per the Data Privacy Act, must be upheld at all times;
  - For further guidance, evaluators and evaluation managers may refer to the Helsinki Declaration, the Nuremberg Code, and other frameworks for ethics in human research.
- Evaluators shall be sensitive to the cultural, social, and economic environment of all stakeholders, and conduct themselves in a manner that is fair, appropriate, and sensitive to this environment.
  - Evaluators shall be accountable for their performance and their products. In addition:
    - They should strive to conduct the entire process fairly and honestly; and present their findings and recommendations in a clear, accurate, and balanced manner
    - In cases where they uncover evidence of wrongdoing, they should report such cases discreetly to the appropriate investigating body. Section 5.3 provides additional guidance on ethical considerations in doing evaluations.



**Transparency:** Access to evaluation information facilitates knowledge sharing among stakeholders, enhances the credibility of both the process and the results, and inspires greater confidence in implementing agencies and the government in whole. The entire process—from setting the Terms of Reference (ToR) and selection of evaluation teams, to the dissemination of key findings—should be made easily understood by all stakeholders and shared in accessible formats (DAC, 2010; WB, 2007; UNEG, 2018).

All final evaluation reports should be public documents (WB,

2007; UNEG, 2018). To ensure transparency in processes, results, and recommendations, the NEPF mandates agencies to:

- Describe the evaluated project/program;
- Provide adequate background and context including the purpose of the evaluation and the evaluation issues and questions;
- Describe the evaluation methodology, including limitations and the approaches adopted to mitigate these;
- Disclose the identities of the principal members of the evaluation team and, if applicable, the evaluation steering committee or other advisory bodies;
- Clearly state evaluation findings, along with a description of the evidence on which each finding is based; and
- Clearly distinguish evaluator opinions (e.g., recommendations) from statements of fact (e.g., findings).



**Human rights and gender equality:** Evaluations should protect, promote, and uphold the universally-recognized values and principles of human rights, and address issues involving gender and underrepresented groups. Each evaluation goal, methodology, finding, and recommendation should embrace and reinforce the “no one left behind” mandate that influences the Philippines’ commitments under the SDGs, as ideally spelled out in its agendas for development (e.g., Ambisyon 2040 and PDP).

The following guideposts should also be considered to ensure a rights-sensitive approach to evaluations:

- The evaluation process should show sensitivity to gender, tradition, and basic rights of all stakeholders. Evaluators should

also be sensitive to inequity, discrimination, and related issues around the social and cultural environments in which they work;

- Evaluation teams must strive to be gender-balanced; and the objectives, as gender-sensitive and empowering as possible;
- In generating and reporting data and information, evaluators should strike a balance between the stakeholders' rights to access information and confidentiality.



**National capacity and ownership:** When used effectively and conducted in adherence to widely-accepted norms and standards, public sector evaluations can build a culture of learning and accountability, and consequently, feed into the improvement of government planning, budgeting, and implementation towards the achievement of country-wide and sector-based development goals. Scaling up the evaluation capacity at the national level entails the following:

- Government support for creating an enabling environment (e.g., legislation of evaluation policies, allocation of resources for evaluation manpower and infrastructure)
- Efforts to build individual competencies, and institutional capacities for commissioning and using evaluations across government agencies; and
- Assistance from development partners, which must be aligned with the needed competencies and capacities of the government, introducing international best practices while ensuring right-fit to the national context.

## 1.7 Applicability of the Guidelines

A. The Guidelines are designed to spell out processes and standards on evaluation across national government agencies. In line with the NEPF, certain programs and projects that meet the prioritization criteria as defined in these Guidelines must be subjected to an evaluation<sup>2</sup>.

For the purpose of these Guidelines, the term “agencies” are used to mean all NGAs, SUCs, GOCCs and GFIs of the Executive Branch. Constitutionally and fiscally autonomous agencies—Congress, the Judiciary, and Constitutional Commissions—as well as GOCCs and GFIs that do not receive budgetary support from the National Government are not required but are encouraged to use these Guidelines.

B. Local government units (LGUs) are not covered by these Guidelines. Similarly, the Regional Offices (ROs) and subnational units of NGAs may refer to these Guidelines for the general principles, standards, and processes of evaluations.

C. A program is defined as:

“[A] group of activities and projects that contribute to a common particular outcome. A program should have the following: (a) unique expected results or outcomes; (b) a clear target population or client group external to the agency; (c) a defined method of intervention to achieve the desired result; and (d) a clear management structure that defines accountabilities” (DBM, 2018; DBM, 2017)

A program may be divided into subprograms, which have either a more specific method of intervention or a more defined set of clients.

A project is defined as a special undertaking carried out within a definite timeframe and intended to result in some predetermined measure of goods and services; while an activity is a recurring work process that contributes to the implementation of a program or subprogram (DBM, 2017).

<sup>2</sup> This current version of the Guidelines does not yet cover evaluations covered by local government. However, the Guidelines will be expanded in future iterations to include such.



The Guidelines does not cover the evaluation of policies, but such may follow the principles and procedures stated herein.

- D. The NEPF mandates that all programs and projects be evaluated at least once, i.e., at the end of their life cycle, and otherwise as frequently as necessary. Hence, a budget proposal for any new program or project should include an allotment for evaluation at least at the end of its implementation.

Evaluations shall be designed sufficiently to provide timely answers to critical questions with an adequate level of certainty, but not costlier than necessary.

Depending on the size and complexity of a program being evaluated, evaluations can range from simple desk reviews to complex studies involving multiple sources of data and various methods.

- E. The sections within the Guidelines will be reviewed and updated periodically to consider arising needs in evaluation guidance, as well as changes in government policies.

## 1.8 Overview of the evaluation process

- A. To deliver findings that are useful enough to inform decision making and improve program delivery, every evaluation process shall, at a minimum:
1. Begin with a systematic evaluability review of the program or project to be evaluated and a consultation with its key stakeholders;
  2. prepare a detailed evaluation plan<sup>3</sup> and ToR that define the objectives, scope, methodology, and financing of evaluation studies;
  3. Require transparency and clarity through reporting of methods, assumptions and key findings
  4. Produce findings that help in informing decisions on program implementation and budget utilization; and
  5. Adhere to quality standards in each stage of the process. (Queensland Treasury and

Trade, 2014).

- B. The evaluation process entails four (4) main phases:

**Initiation:** Identifying an evaluation agenda, checking the evaluability or readiness for evaluation, and defining the scope and requirements through an evaluation plan;

**Preparation:** Agreeing on the management structure of the evaluation, and roles and responsibilities, including the establishment of an evaluation reference group, developing the evaluation ToR, procuring the evaluation consultant/s;



**Implementation:** Briefing and setting expectations with the evaluation team, reviewing the inception report prepared by the team through an evaluation reference group meeting, engaging the ERG to review the draft evaluation report, finalization of the report; and



**Utilization:** Preparing the management response and implement follow-up actions, preparing and disseminating evaluation products and organizing knowledge sharing events, and preparing for the design of future evaluations for the program or project.

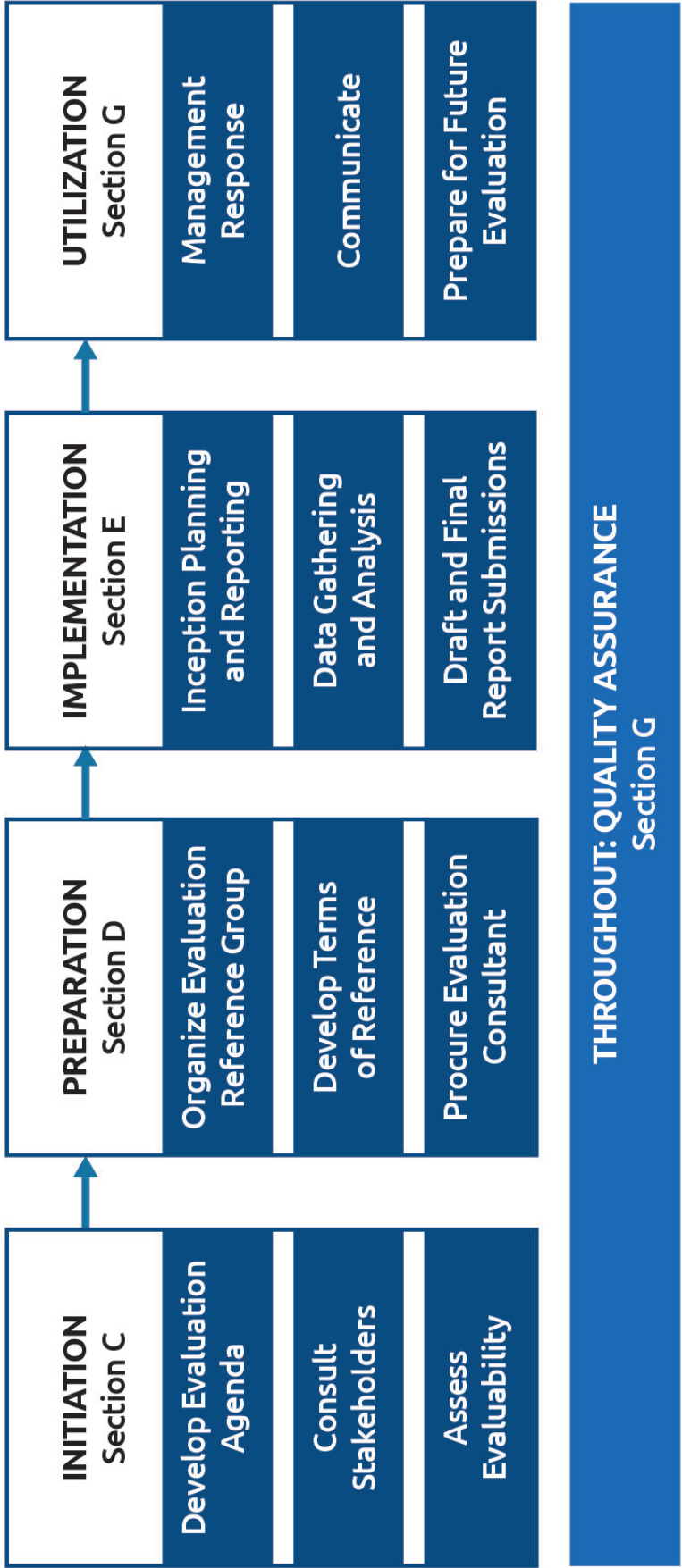


- C. Although performing the steps in the prescribed order is advisable, flexibility is also necessary as the situation evolves during the evaluation process. To ensure evaluations are effectively implemented, some steps may need to be revisited (e.g. scoping and specifications in the event of a failed bid or inadequate data).



<sup>3</sup> See section 3 for detailed guidance on preparing an evaluation plan.

Figure 5: Evaluation phases



Source: UNDP (2019)

Annex A presents an illustrative timeline for the evaluation phases, with the outputs scheduled in time for the budget preparation process.



# INSTITUTIONAL RESPONSIBILITIES

To promote the purposive conduct of evaluations in the public sector, the NEPF assigns responsibilities to NGAs as they undertake the evaluation of their programs and projects. Moreover, it creates an inter-agency Evaluation Task Force (ETF) and its Secretariat to operationalize the NEPF.

In this Version 1.0 of the Guidelines, this section on Institutional Responsibilities is largely a reiteration of the responsibilities assigned by the NEPF to NGAs (NEPF section 6.2) and the ETF and its Secretariat (section 6.2). Further guidance will be issued to thresh out the roles and responsibilities, including: governance arrangements within NGAs for undertaking evaluation; organization and composition of agency evaluation units; and staffing and competencies.

## 2.1 Implementing agencies

NEPF section 6.1 mandates NGAs to: (i) prepare rolling six-year evaluation agendas; (ii) submit evaluation plans with all their program or project proposals; (iii) provide adequate management response to evaluation findings; and (iv) use evaluation results for planning and programming. It also mandates agencies to form evaluation units to support it in undertaking these functions.



**Evaluation Agenda** – Agencies shall formulate and maintain a rolling six-year evaluation agenda to coincide with the timeframe of the PDP and PIP (refer to section 3.1 of these Guidelines).



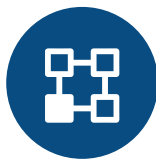
**Evaluation Plan** – Agencies shall ensure that all the programs or project proposals that they put forward for annual budgeting are submitted with an evaluation plan (refer to section 3.2 of these Guidelines).



**Management Response** – Agencies shall craft an appropriate management response to evaluation findings and recommendations, identify follow-through actions to such findings, and specify the time frame to implement the proposed actions (refer to section 6.1 of these Guidelines).



**Use of Evaluations for Planning** – Agencies shall ensure that evaluation results are used as inputs to planning and budgeting processes. In particular, they shall consider the results of previous evaluations in designing proposals for new programs and projects, and in modifying the existing ones (refer to sections 1.4 and 1.5 of these Guidelines for how evaluations should link with the planning, investment programming, and annual budgeting processes).



**Formation of Evaluation Units** – Agency heads are responsible for establishing capable and neutral evaluation units at the central office level and reporting directly to the agency head. This is subject to existing DBM policies, rules, and regulations on organizational and staffing pattern changes (further guidelines will be issued in the future). Evaluation units shall:

- Submit NGA's rolling six-year evaluation agenda to agency head for approval;

- Prepare an evaluation plan for inclusion in all funding requests for budgetary support;
- Oversee implementation of evaluations and ensure they are undertaken impartially and in line with the provisions of these guidelines;
- Manage the agency's evaluation budget and related activities;
- Submit to the agency head the findings and recommendations of evaluation activities;
- Coordinate the preparation of a Management Response to evaluation findings and recommendations, including follow through actions by concerned units;  
Ensure timely publication on agency public website of all evaluation reports;
- Submit evaluation reports to the Evaluation Task Force; and
- Serve as a repository for all agency-level evaluations carried out.

Future guidelines on evaluation units will be issued. In the meantime, agencies may designate existing planning or M&E units to perform the above.

## 2.2 Evaluation Task Force and Secretariat

Oversight of the NEPF is the responsibility of the Evaluation Task Force (ETF). The following reiterate pertinent points of section 6.1 of the NEPF:

A. Members of the ETF are:

- Secretary, National Economic and Development Authority (Chair)
- Secretary, Department of Budget and Management (co-Chair)
- Office of the President-Presidential Management Staff (OP-PMS)
- The ETF may request other government agencies (voting) and invite civil society, academe, and private sector (special non-voting) to participate.

B. The functions of the Evaluation Task Force are to:

- Provide overall policy direction and coordination on the evaluation agenda of the public sector
- Report to NEDA Board on all evaluations conducted in the public sector
- Authorize and commission the conduct of evaluations on top of those performed by the implementing agencies
- Issue evaluation standards and guidelines
- Review and provide feedback on the evaluation agendas of Agencies
- Adopt a sanctions and incentives system
- Ensure the creation of appropriate institutional structures to mainstream the NEPF.

C. The Evaluation Task Force:

- May authorize the creation of a sub-cabinet level Technical Committee composed of NEDA, DBM, OP-PMS and PSA. The Technical Committee may be tasked, among others, to review the quality of evaluation reports;
- May also consider the creation of Regional-level evaluation task forces;
- Shall link with existing NEDA Board Committees and other relevant Committees (e.g., DBCC) to ensure policy coherence.

D. The ETF Secretariat provides technical and administrative support to the task force.

Currently the NEDA Monitoring and Evaluation Staff (MES) serves as the ETF Secretariat in the interim. Its responsibilities include the following:

- Recommend to the ETF policies, strategies, and guidelines for the effective implementation of the NEPF
- Prescribe the format and content of an evaluation plan;
- Monitor and report on progress and results of evaluation activities undertaken by Agencies;
- Conduct capacity development activities jointly with NEDA and DBM for the operationalization of the NEPF;
- Conduct/manage evaluations as authorized by the ETF;

- Provide Secretariat support to the ETF;
- Recommend sanctions and incentives;
- Formulate criteria for evaluations to be endorsed for ETF approval;
- Prepare a consolidated report of individual evaluations for consideration of the ETF (and/or the Technical Committee).

## 2.3 Role of Regional Offices

A. Although the NEPF did not explicitly state their responsibilities, the regional offices of NGAs play an important role in evaluation insofar as providing the regional development perspective.

B. On Agency Regional Offices. Agencies' evaluations are driven from the central office, particularly by their respective evaluation units. Still, agency evaluation units are advised to consult their regional offices in:

- The determination of programs and projects to be included in the agency's evaluation agenda;
- The design and conduct of evaluations; and
- In the formulation of management responses to evaluation findings and recommendations, especially those which pertain to or are to be implemented by the regional offices.

C. On NEDA Regional Offices (NROs). NROs play a crucial role in regional development planning, monitoring, and evaluation. For one, the NEDA Regional Directors serve as Vice Chairpersons of the Regional Development Council (RDCs) and Chairpersons of the Regional Project Monitoring Committees (RPMCs). The NROs may refer to these Guidelines for general principles, standards, and processes in conducting evaluations. Particularly, NROs are encouraged to develop evaluation agendas and align them with their respective Regional Development Plans (RDPs), consistent with the M&E work programs of their RPMCs, and in consultation with relevant AROs and other members of RDC and RPMC. Evaluation findings and results must also be presented before the RDC and/or RPMC for appropriate action by the members.



# INITIATING AN EVALUATION

## 3.1 Developing an agency evaluation agenda

- A. The NEPF mandates agencies to formulate and maintain a rolling six-year evaluation agenda, to coincide with the timeframe of the PDP and PIP. The multi-year agency Evaluation Agenda (see Template 1), which is to be developed during the first year of the plan, lists the programs and projects that will be evaluated throughout the PDP and PIP period. The approved agency evaluation agenda may be periodically updated—annually, for instance—to include any new priorities identified under the PDP Midterm Update and the annual Budget Priorities Framework.
- B. Several parameters should be considered when identifying programs and projects that should be prioritized for inclusion in an agency’s evaluation agenda. The agency may select programs and projects to evaluate if they meet at least one or more of the parameters identified.
- C. Each evaluation listed in the evaluation agenda should have the following information:
  - Name of agency unit commissioning the evaluation
  - Name of the program or project to be evaluated
  - Identification of relevant PDP and PIP Chapter(s) and SDG Target(s)
  - Total budget of the program or project and its multi-year allocation
  - Status: new, ongoing

## 3.2 Developing an evaluation plan

- A. For all proposed programs or projects, an Evaluation Plan (see Template 2) shall be prepared in accordance with best practices (NEDA & DBM, 2015). The evaluation plan shall be submitted, along with a comprehensive project profile or feasibility study, as a basis for funding.
- B. An evaluation plan identifies key evaluation questions and maps out the strategic evaluation methodology for a particular program. It incorporates the overall purpose for the evaluation, including the key evaluation projects and reports across the programs’ life cycle, as well as activities related to the evaluation, such as data collection techniques unique to the particular evaluation and the process used to build capacity for evaluation. Planning likewise entails choosing the most strategic type of evaluation—one that provides substantive information to support the right decisions. Hence, it is important to figure out which questions need to be answered, how they will be answered, and by when (Queensland Treasury and Trade, 2014; UNDP, 2019).

- C. Each evaluation plan should respond to the following basic questions (Queensland Treasury and Trade, 2014; UNDP, 2019):

**What:** the scope, objectives, key questions, and other elements that should guide the evaluation, including a results framework, theory of change, or project logic map;

**Who:** those responsible for managing, for carrying out, and for providing the information required for the evaluation;

**How:** the methods of gathering and analyzing data to answer evaluation questions, as well as strategies for risk management, and the reporting and use of results;

**When:** the specific period/s of performing the evaluation within the program cycle (e.g., aligning the process with critical phases of budgeting and legislation), key milestones, and deliverables;

**Resources:** the people, materials, infrastructure, logistics, and cost needed for the evaluation.



- D. The scale, scope, and cost of a program determine the depth and complexity of an evaluation plan. Evaluation plans shall take into consideration results of previous evaluations of similar programs and refer to relevant evaluation findings, recommendations, and resulting changes to the proposed program. In cases where recommendations were not followed, the proposal shall include an explanation. For existing programs, program managers shall review whether evaluation arrangements are in place and if not plan to ensure an evaluation plan is developed and implemented.
- E. Planning for an evaluation shall begin right when the program and its implementation strategies are being designed to ensure the collection of data that are necessary for coming up with useful evaluation findings (e.g., for comparison between baseline, mid-term, and completion). Ongoing collection and monitoring of data also facilitate continuous program improvement.
- F. Developing an evaluation plan is an iterative process, where the document may be updated several times before proceeding with the evaluation itself. An initial, incomplete version may be prepared with the evaluation agenda. The plan may then be improved and completed as the program's readiness for evaluation is assessed and requirements for evaluability are met (see below). Ideally, the evaluation plan is attached to the ToR of the evaluation consultants to be hired.

### 3.3 Assessing evaluability

- A. An evaluation requires several conditions to ensure a favorable outcome. For one, baseline data, data collection systems, and results frameworks are important prerequisites for projects to be evaluated in a reliable and credible fashion (OECD-DAC 2010). Moreover, the degree of support and demand for the evaluation—in terms of not only the availability of resources but also of the broader socio-political context—could spell its success or failure.
- B. An Evaluability Assessment (included in Template 2 – Evaluation Plan) examines the extent to which a program or project is ready to be evaluated. It will also help in clarifying the objectives for the evaluation, identifying the best methodology given the available data and resources, and planning for external factors that could influence the success

or failure of the evaluation. It can also help the agency decide to instead undertake an internal assessment in case of failure to meet basic evaluability requirements.

- C. An evaluability assessment addresses the following issues (World Bank, 2016; and UNDP, 2018):



**Clarity of the intervention** – does the subject of evaluation have a clear logical framework or theory of change? Are the objectives, outcomes, and outputs clearly defined? Are the indicators clearly stated?



**Availability of data** – is sufficient data collected against the indicators? Is there baseline data? What methodology can be used given the available data?



**Stakeholder interest and intended use** – how can decision-makers use the evaluation to improve program design, implementation, and resource allocation? Are there socio-political factors that could hinder the conduct of the evaluation?



**Availability of resources for the evaluation** – are there enough financial, human, and knowledge resources to conduct the evaluation? How much is required?

- D. Below is a detailed evaluability checklist that evaluation managers may use in assessing the readiness of the program or project for evaluation. This has been integrated in the evaluation plan template for ease of reference:

- Generally, an evaluation can proceed if “yes” is the answer to majority of the questions above. Nevertheless, the evaluation manager must proceed with caution and work towards addressing the issues and gaps identified.
- If a program or project is deemed not yet ready for an evaluation (i.e., less than majority are “yes” responses to the questions above), the evaluation manager, in partnership with the program or project manager, must undertake some work that would strengthen the evaluability of the program or project. These include:

- 1. Developing a results framework or theory of change (TOC) for the program or project, accompanied by a costed M&E plan;
- 2. Developing a data collection system, collecting baseline data or undertaking a “midline” data collection in case of an ongoing program or project;
- 3. Engaging policymakers and stakeholders to support the evaluation especially in allocating resources, mitigating risks, and creating demand;

Table 2. Evaluability Checklist

Adapted from: UNDP Evaluation Guidelines (2018 Update)

		YES	NO
Results Framework			
1	Is the Results Framework and/or Theory of Change clearly defined?	<input type="checkbox"/>	<input type="checkbox"/>
2	Are the indicators clearly stated in the results framework?	<input type="checkbox"/>	<input type="checkbox"/>
Key evaluation questions			
3	Is the planned evaluation relevant and useful to key stakeholders?	<input type="checkbox"/>	<input type="checkbox"/>
4	Are the stakeholders committed to support the evaluation?	<input type="checkbox"/>	<input type="checkbox"/>
5	Are the evaluation questions feasible given the: (1) project design, (2) data availability, and (3) resources available?	<input type="checkbox"/>	<input type="checkbox"/>
6	Are the evaluation questions of interest to key stakeholders?	<input type="checkbox"/>	<input type="checkbox"/>
Data			
7	Is there sufficient data collected to answer the evaluation questions? Was such data collected at baseline and consistently collected at various intervals?	<input type="checkbox"/>	<input type="checkbox"/>

		YES	NO
8	Is there sufficient data disaggregation (e.g. age, sex, disability, ethnicity, migratory status, and geographic location where relevant)?	<input type="checkbox"/>	<input type="checkbox"/>
9	If data, particularly baseline data, is not available, are there plans and/or means available to collect and disaggregate the data?	<input type="checkbox"/>	<input type="checkbox"/>

Risk Identification and Analysis

10	Will physical, political, social, economic, and organizational factors allow for an effective conduct and use of evaluation as envisaged?	<input type="checkbox"/>	<input type="checkbox"/>
----	---	--------------------------	--------------------------

Evaluation Timeline

11	Is there sufficient time for the evaluation?	<input type="checkbox"/>	<input type="checkbox"/>
----	--	--------------------------	--------------------------

Evaluation Timeline

12	Are there available service providers?	<input type="checkbox"/>	<input type="checkbox"/>
13	Are there sufficient human resources?	<input type="checkbox"/>	<input type="checkbox"/>
14	Are there sufficient financial resources?	<input type="checkbox"/>	<input type="checkbox"/>

3.4 Selecting the evaluator

A. In all situations, evaluations must be conducted in an independent and transparent manner. To ensure impartiality, ideally, all evaluations should be externally conducted, whether commissioned to an individual consultant, a firm, or an independent research entity (e.g. PIDS, HEIs).

B. In certain instances, given an agency’s resources, individual program characteristics, and evaluation capacity, an internal evaluation may be considered. An agency’s staff may be



able to perform evaluation if the highest possible degree of impartiality can be ensured to maximize the objectivity and credibility of the results and minimize bias and conflict of interest.

An evaluator that is internal to the organization may be an individual or unit reporting directly to the head of agency or governing board ideally; or at the very least not belonging to the program or project design and implementation unit and/or not reporting to an agency official who is responsible for leading the implementation of the program or project being evaluated. If the conditions above are not met, the exercise is considered an internal assessment that need not strictly follow these Guidelines.

## 3.5 Stakeholder engagement

A. The success of a program evaluation lies in the identification of and engagement with key program stakeholders to clarify its purpose. The consideration of stakeholder engagement (Preskil and Jones, 2009), including fostering input, participation, and power sharing among those with an investment in the conduct of an evaluation and its findings, is an important first step when planning an evaluation.

If stakeholders are involved in the planning, design, conduct, and follow-up stages of an evaluation (UNEG, 2018), not only will it help to identify what they want to know, they will more likely commit to the evaluation and use the results of the evaluation for decision-making.

B. Evaluation stakeholders typically fall into three major groups:

- Those directly involved in running the program or project (e.g., program/project managers and staff, funding partners, and delivery units).
- Those served or impacted by the program (e.g., program participants and their families, individuals who avail of or benefit from the program, and the general public).
- Those who are interested in the program and its intended impact, and would use the

evaluation results for policymaking and lobbying, among others (e.g., senior public and development sector managers and elected officials; as well as community, private sector, or civil society groups) (Queensland Treasury and Trade, 2014; UNDP, 2019).

C. Identifying and considering the needs of key stakeholders can inform the evaluators on key process and decision points, including:

- How the evaluation can be best fit into decision-making cycles to inform policy and program improvement
- How the evaluation can be designed, presented, and disseminated for relevance, credibility, acceptability, and overall effectiveness
- How evaluation findings can be robust and comprehensive enough to be to support the intended end use and meet the level of scrutiny the findings will be subject to (Queensland Treasury and Trade, 2014; UNEG, 2016).

D. In challenging cases where multiple agencies or proponents are working together to deliver a program, the management of stakeholder expectations will help ensure evaluation objectives are met and kept relevant, and the findings can be used to inform decision-making (Queensland Treasury and Trade, 2014).



# PREPARING AN EVALUATION

## 4.1 Program objectives and intended outcomes

- A. Evaluation design entails a complete understanding of how a program is intended to work, what it tries to achieve in terms of measurable objectives, and the underlying policy problem it seeks to address. Even the most rigorous evaluation will fail if it is not clear what impact the program seeks to demonstrate and why. (Queensland Treasury and Trade, 2014).
- B. Evaluations are most effective when program objectives are clearly defined, and success indicators are built around and measured against these objectives.
- C. Evaluators and program managers should identify the following elements to gain a complete understanding of the program:



### Need

Why the program is required



### Objectives

What the program aims to achieve and why



### Inputs

The resources needed to operate the program (e.g., labor, materials)



### Short-run outcomes

Changes in beneficiaries' awareness, knowledge, skills, and attitude



### Outputs

Direct products of program activities (e.g., types of services to be delivered)



### Activities

Processes, tools, events, technology, and actions integral to program implementation



### Medium-run outcomes

Changes in behavior



### Long-run outcomes

Wider economic, environmental, and social impacts of a program.

These elements are typically included in a results framework(4, SUPERScript PLS) that is required for every program that receives budgetary support. The original results framework, with any necessary revisions to reflect updates in program design, needs to be carried across and incorporated into the evaluation plan.

# 4.2 Evaluation objectives

- A. Evaluation approaches may range from simple desk reviews to complex studies involving multiple sources of data. In some cases, a formative evaluation is deemed appropriate midway through project or program implementation, in addition to the summative evaluation at the end of its life cycle or of the PDP period.

The scale and cost of every evaluation typically expands in relation to the following factors: the level of ambiguity of program or project results; the potential consequences of a program or project; its size, complexity, and uniqueness; and the information needs of managers and stakeholders (see Annex D of NEPF Guidelines).

- B. Thus, it is important to clearly set the objectives of an evaluation to ensure it produces the information stakeholders need to take an informed action.

For example, an assessment of efficiency may inform stakeholders on the cost-effectiveness of an intervention, but this might be irrelevant if the intervention lacks enough relevance, appropriateness, and effectiveness to meet the needs of a client, a stakeholder, or the broader community (Queensland Treasury and Trade, 2014).

Many evaluations involve assessments of multiple elements of a program (e.g., effectiveness and efficiency) to ensure stakeholders and decision makers get a more complete picture of the program, but it is important to find a balance between the desired level of information and the resources required to produce insightful evaluation findings.

- C. Evaluations shall consider the following questions covering the widely- recognized criteria (OECD-DAC, 2002; NEDA & DBM, 2015; OECD-DAC, 2019; UNDP 2019) for evaluation (See Annex D for more sample evaluation questions):

<sup>4</sup> For more details on results frameworks see Section 3 of NEDA 2017. Results-based monitoring and evaluation (RbME) guidebook. Draft. Manila: National Economic and Development Authority. The results framework will also be reflected in the Evaluation Plan (Template 2).



## Relevance

- A. **Responsiveness to stakeholder needs.** To what extent does the program address the urgent needs of the stakeholders? Did the project design and choice of outputs and activities properly reflect and address the needs of communities?
- B. **Programmatic alternatives.** Are there better ways to achieve the outcomes/impacts of programs/projects, or for programs/projects to contribute to related national priorities?
- C. **Project design.** How valid is the Theory of Change? Were the planned and actual outputs and activities of the project consistent with the intended outcomes?



## Coherence

- A. **Compatibility with other interventions, may be Internal or External.** Do other interventions and policies support or undermine the program/project being evaluated, and vice versa?
- B. **Internal coherence.** To what extent is the program/project harmonized with the other interventions of the government? To what extent do program outcomes/impacts align with the achievement of national priorities and existing laws, including PDP sector outcomes? How consistent is the intervention with the relevant international norms and standards adhered to by the government?
- C. **External coherence.** How consistent is the program/project with the interventions of other actors in the same context? How relevant is the

program/project with the international norms and standards adhered to by the government? How does it complement, harmonize, and coordinate with such other actors' interventions? Does the intervention add value without duplicating effort?



## Effectiveness

- A. **Achievement of objectives.** What intended outputs and outcomes/impacts (short-term outcomes, medium-term outcomes, and long-term impacts) were found, and to what extent can they be attributed to program activities? How likely will the project/program contribute to the planned outcomes?
- B. **Unintended results.** What unintended outputs and outcomes/impacts were found, and to what extent can they be attributed to program activities?
- C. **Coverage.** Is the project reaching the intended beneficiaries, rights holders, and duty bearers?
- D. **Value added.** What value has the implementing agency or the project/program added?



## Efficiency

- A. **Efficient delivery of outputs.** Were the activities cost-efficient? Was the utilization of resources optimized in terms of the realization of the program objective? To what extent was resource utilization minimized in relation to the delivery of outputs? Are the variances between planned and actual expenditures justified?

B. **Operational alternatives.** Are there better, more efficient ways to deliver program outputs?

C. **Timeliness.** Were the objectives achieved on time?



## Sustainability

- A. **Continuity of benefits.** How likely will the benefits of the project continue in the future? To what extent did the benefits of a program continue after funding ceased?
- B. **Success factors.** What were the major factors which influenced the achievement or non-achievement or sustainability of the program?



## Impact

- A. **Overall effects of intervention vis-à-vis counterfactual.** What changes or effects has the program/project brought? What would have happened if the intervention had not taken place?
- B. **Tangible effects vis-à-vis baselines.** What real difference has the intervention made to the intended beneficiaries? How do the changes fare from the initial state/situation of the beneficiaries?
- C. **Effects to stakeholders.** How many people have been directly and indirectly affected by the interventions?



## Other Criteria

- A. In addition to the “classical” criteria above, evaluations may also make use of other evaluation criteria depending on the needs of the sector of the program or project to be evaluated or the issues around it. Some of these cross-cutting criteria include Gender, Human Rights, and Peace (see Annex D for sample evaluation questions on these criteria).
- D. Specific evaluation questions are developed from the program’s results framework and should be grouped according to the criteria of relevance, coherence, effectiveness, efficiency, and sustainability.
- E. Evaluation questions define the information that the evaluation should provide, focus the evaluation, and drive decisions on appropriate data collection and analysis while also helping to determine the type of evaluation required (Queensland Treasury and Trade, 2014; UNDP, 2019).
- F. Guide evaluation questions should be outlined in the evaluation plan and in the resulting ToR. They should be further refined by the evaluation team and agreed with by the members of the evaluation reference group during the inception phase.

## 4.3 Evaluation methods

- A. The evaluation questions chosen by stakeholders define the appropriate methodologies to be used for the evaluation. Conversely, the evaluation type and methods that can be used will depend on the evaluation objective and questions.
- B. After deciding on the objectives of the evaluation and the key questions that the evaluation will need to address, the evaluator will need to consider:



**Methods:** Which methods are going to be used to answer the evaluation questions? The choice of method(s) will depend on what information is needed, the program characteristics, and organizational capacity and resources. Evaluation approaches can be grouped into the following broad categories:

- **Formative.** These evaluations provide information on how a program may be developed (new programs) or improved (both new and existing programs).
- **Process.** These evaluations measure what is achieved by the program, investigates how the program is delivered, and can be used to improve programs by informing adjustments to delivery.
- **Summative.** These are conducted at the end of an intervention to determine the extent to which the anticipated outcomes were produced.

Table 3 illustrates the types of questions that can be addressed by different evaluation approaches. See Annex B for more detailed information and discussion. on evaluation approaches and methods.

Table 3: Key evaluation questions by type of evaluation

Type	Description	Example questions to be
Formative	<ul style="list-style-type: none"><li>■ Formative evaluations provide information on how a program may be developed (new programs) or improved (both new and existing programs). Examples include:<ul style="list-style-type: none"><li>● Needs assessment - to determine who needs the program, the scale of need, and what may meet the need.</li><li>● Program logic map - to ensure a clear picture of how and why the program will produce expected results. Business case: To define the program, delivery methods, the target population, and the possible results.</li><li>● Cost-benefit analysis - to compare the cost of the program with the value of most of the program's benefits. This allows for program decisions to be analyzed.</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ What is the problem to be solved?</li><li>■ What are the characteristics and needs of the target population?</li><li>■ Where is the problem?</li><li>■ What is the most appropriate plan of action to address the problem?</li><li>■ Is government intervention appropriate?</li><li>■ How would we measure/recognize success?</li><li>■ Is the project's Theory of Change sufficient?</li></ul>
Process	<ul style="list-style-type: none"><li>■ Process evaluations help in differentiating ineffective programs from failures of implementation. Alternative delivery solutions may be</li></ul>	<ul style="list-style-type: none"><li>■ How is the program being implemented?</li></ul>

	<p>explored and investigated. Examples include:</p> <ul style="list-style-type: none"><li>● Collection of descriptive statistics – e.g., age, race, marital status, education, income, and number of children, to identify who benefits most from the program.</li><li>● Process outcomes - a description of the status or condition of participants after they participate in the program.</li><li>● Process evaluation tools - include: process logs, attendance sheets, mailing list, telephone call log, and participant satisfaction survey.</li><li>● Case study - a method for developing a complete understanding of a process, program, event, or activity. A common element of a case study is systematic and detailed data collection from multiple sources, particularly first-hand experiences.</li></ul>	<ul style="list-style-type: none"><li>■ Are the activities being delivered as intended?</li><li>■ Are program participants being reached as intended?</li><li>■ Could activities be delivered by an alternative method or means?</li></ul>
Summative	<ul style="list-style-type: none"><li>■ Summative evaluation reports when the program has been running long enough to produce results. Examples include:<ul style="list-style-type: none"><li>● Outcome evaluation - determines whether the program has caused the</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ What are the net effects of the program?</li><li>■ Is the program achieving/has it</li></ul>



	<p>intended effect on the target results.</p> <ul style="list-style-type: none"><li>● Impact evaluation – determines the broader, longer-term changes that occurred as a result of the program.</li><li>● Cost effectiveness analysis - relates the costs of a program to its key results or benefits. It is a tool that is often used for health programs.</li><li>● Cost-benefit analysis - allows for program decisions to be analyzed. Both cost effectiveness and cost-benefit analysis address the questions of efficiency by standardizing results in terms of their dollar value to answer questions of value for money.</li></ul>	<p>achieved the intended results?</p> <ul style="list-style-type: none"><li>■ o what extent can changes be attributed to the program?</li><li>■ Is the program the best use of the resources that it costs?</li><li>■ What would be the impact of cessation?</li><li>■ What alternative service delivery options, if any, can be considered?</li></ul>
--	--	--

Source: NEDA Ex-Post Evaluation Manual (2014); Western Australia Government (2015).



**Data collection and analysis:** What information is required and from whom and how the information can best be obtained? The evaluation plan should spell out the data points can best measure program relevance, coherence, efficiency, effectiveness, impact, and sustainability, and the most functional method for sourcing or collecting those. Ideally, data collection for evaluations should be a mix of qualitative and quantitative methods.

The plan should also identify and consider any cultural, ethical, privacy concerns that may affect the collection and use of data, as well as

strategies to address any limitations or deficiencies in data collection. (See Annex C for a discussion of different methods for collecting and analyzing data, and the issues to consider in selecting methods.) Following the Philippine Statistical Act of 2013 (Republic Act 10625, the design and instruments of statistical surveys or censuses sponsored and/or to be conducted by government agencies including government corporations at the national and/or sub-national level are subjected to the Philippine Statistics Authority (PSA) Statistical Survey Review and Clearance System.



**Reporting:** How is the information going to be used? Requirements for analyzing and reporting evaluation data will depend on both the evaluation objectives and how the intended audience can make the best use of its findings. The scope of analysis should allow for the development of valid findings, as well as the best mechanism for communicating them to stakeholder groups, e.g., the format, language and structure of reporting.



# 4.4 Risk identification and assessment

A. A risk is defined as any uncertain event that may bear a positive or negative impact on a program and its evaluation. Unexpected events can affect a program’s implementation and, as a consequence, its evaluation. Risk identification allows the program evaluation team to proactively approach potential or real problems, rather than defensively respond to problems after they occur. Risks may be categorized under broad headings (see Table 4).

Table 4: Risk Categories and Identification Questions

Risk Category	Example questions to determine risk factors
Approvals	Are there risks to securing the approval of relevant officials or offices (e.g., funding institutions) for the evaluation?
Design & scope	Is the program/project to be evaluated clearly defined? Are adequate data collection and analysis tools available?
Procurement	Is there enough supply market of evaluation consultants (individuals / firms) and other specialists required?
Budget and funding	Has there been a budget set aside for the evaluation? Are there enough funds to complete the evaluation?
Inputs and cost	Are there risks of cost escalation due to the market price of consultants, technologies, and other inputs to the evaluation project?
Consultant performance	Will evaluation team members and stakeholders be available when needed? Are there risks to their performance?
Agency capacity	Does the agency have enough manpower to manage the evaluation?
Institutional support	Do all stakeholders agree on the evaluation’s objectives and purpose? Are there any risks posed by withdrawal of support to the evaluation?

Legal and policy	Are there any regulatory or ethical issues that may impact the evaluation?
Sustainability	Will the evaluation be useful for future programming?
Force Majeure	Are evaluators likely to be put in danger while conducting the evaluation? Are there force majeure events that could put the evaluation to risk?
Other	For other risk classifications, please specify.

Source: NEDA MES Risk Analysis Guide (n.d.)

- B. It may also be useful to consult relevant historical information from previous similar program evaluations that may include lessons learned which describe problems and their resolution. Once potential risks are identified in the evaluation plan, it is useful to develop measures to prevent or minimize their occurrence.
- C. To minimize risks, evaluation managers will benefit from a contingency plan that spells out appropriate responses to emerging problems, and measures to ensure the evaluation stays on track to meet its objectives (see Table 5).

Table 5: Risks and risk management strategies

Risk Category	Possible risk management strategies
Approvals and Institutional Support	<ul style="list-style-type: none"><li>■ Ensure that the need for evaluation is clearly communicated to senior officers and to funding institutions.</li><li>■ Communicate regularly with government stakeholders to ensure their needs are being met.</li></ul>

Risk Category	Possible risk management strategies
	<ul style="list-style-type: none"> <li>■ Ensure stakeholders know the time involved in carrying out different</li> <li>■ Evaluation approaches.</li> </ul>
<b>Design &amp; scope</b>	<ul style="list-style-type: none"> <li>■ Identify and clearly communicate data requirements at the onset of program planning to allow for the collection of needed data during program implementation.</li> <li>■ Attempt to access proxy indicator from an alternative credible source in cases where data is unavailable for the most suitable indicator.</li> <li>■ Modify research design when appropriate data cannot be sourced.</li> </ul>
<b>Procurement</b>	<ul style="list-style-type: none"> <li>■ Conduct market research on the evaluation capacities required and seek out partnerships, e.g., with state universities and colleges</li> </ul>
<b>Funding and costs</b>	<ul style="list-style-type: none"> <li>■ Allocate funds for evaluation as a separate item in the program or project budget.</li> <li>■ Ensure budget proposals are drafted in line with established government processes.</li> <li>■ Ensure that the evaluation objectives are appropriately communicated, and that findings can be used to inform budgeting decisions.</li> </ul>
<b>Consultant performance and agency capacity</b>	<ul style="list-style-type: none"> <li>■ Evaluators and evaluation managers should regularly monitor milestones vis-a-vis the evaluation plan.</li> <li>■ Give ample time in between evaluation milestones to allow for a small amount of project creep.</li> </ul>

<b>Legal and policy</b>	<ul style="list-style-type: none"> <li>■ Incorporate legal, policy, and ethical issues in the design of evaluation</li> </ul>
<b>Sustainability</b>	<ul style="list-style-type: none"> <li>■ Evaluators should regularly monitor milestones vis-a-vis the evaluation plan.</li> </ul>

Source: NEDA MES Risk Analysis Guide (n.d.); Queensland Treasury and Trade (2014);

## 4.5 Financial and human resource requirements

A. An evaluation plan shall include a budget that itemizes the financial costs (see Table 6) together with an estimate of the human resource requirements. Likewise, the budget shall be accompanied by a timeline that details evaluation milestones, deliverables, and modes for reporting on progress towards the completion of the final evaluation report (see Template 4: Costing an Evaluation).

Table 6: Typical cost components of an evaluation

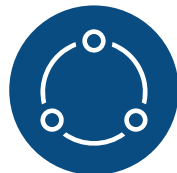
Source: UNDP (2019)

Risk Category	Possible risk management strategies
Approvals and Institutional Support	<ul style="list-style-type: none"> <li>■ Professional fees</li> <li>■ Travel costs (International and domestic)</li> <li>■ Per Diem costs (time in the field)</li> <li>■ Insurance</li> </ul>
Evaluation Implementation and Data Collection costs	<ul style="list-style-type: none"> <li>■ Focus Group and workshop related costs</li> <li>■ Survey costs</li> <li>■ Travel costs</li> <li>■ Translation</li> </ul>
Evaluation distribution costs	<ul style="list-style-type: none"> <li>■ Reproduction</li> <li>■ Communications</li> <li>■ Dissemination forum</li> </ul>

B. Resource considerations for evaluations include the following:



**Scope**  
Geography – does the evaluation cover a specific locality/ies or a nationwide coverage?



**Coverage** – number and type of evaluation questions. How simple or complex are the questions to be addressed? How comprehensive should the evaluation be?



**Duration** – evaluations could be as short as three months (e.g., rapid project evaluation) or as long as three years (e.g., a multi-stage impact evaluation)



**Approach**

Data Collection Requirements – sample size, level of disaggregation, length of survey questionnaire, etc. will influence the cost. Good quality monitoring data will reduce the cost requirement for primary data collection for the evaluation.

Methodology – a multi-year experimental evaluation (e.g., randomized control trial) will cost more than a process evaluation.



**Risk** – the probability and degree of security, operational, environmental, and other risks influences cost. Risk mitigation measures such as insurance and security arrangements should be part of the evaluation budget.

C. All these considerations in planning for an evaluation are encapsulated in the ToR (see Template 5: Terms of Reference) for the evaluation firm or individual(s) to be procured or academic or research institution to be contracted. Procurement of the evaluator shall be guided by the Government Procurement Reform Act and its Implementing Rules and Regulations (see also Annex D for a detailed guide on procuring evaluations).



# IMPLEMENTING EVALUATIONS

- A. Implementing an evaluation requires clarity on who shall take on the oversight and facilitation of the evaluation process, from initiation to the delivery and well into the quality assurance of evaluation products.
- B. It is also important to consider how the findings and recommendations will be communicated to stakeholders, how the evaluation results will be used to inform future program design and government strategies, and who shall have responsibility for monitoring and ensuring agreed recommendations are implemented in practice.

## 5.1 Evaluation governance arrangements

- A. All evaluations should have a clearly defined management structure or governance architecture as well as established and communicated roles and responsibilities (See Table 7).
- Agencies are given flexibility in identifying staff and offices responsible for fulfilling the following functions based on their respective composition or structure.

Table 7. Evaluation key roles and responsibilities

Person/Unit	Roles and Responsibilities
Evaluation Oversight	<ul style="list-style-type: none"><li>■ Sets the agency’s evaluation agenda, including its evaluation budget, to ensure that the prioritized studies are relevant to the needs and priorities of the agency.</li></ul>
Evaluation Commissioner	<ul style="list-style-type: none"><li>■ Identifies the evaluation objectives and key questions.</li><li>■ Supports the evaluation manager in finalizing the evaluation plan and TOR.</li><li>■ Provides the evaluation manager and the evaluators with access to data and key informants.</li><li>■ Participating as a key informant or in focused group discussions.</li><li>■ Proposes the members of the Evaluation Reference Group. The Evaluation Commissioner and Evaluation Manager may serve as Co-Chairs of the ERG.</li><li>■ Participates in the review of key evaluation products— inception report and draft report—especially with respect to the accuracy of findings and usefulness of recommendations.</li><li>■ Prepares and finalizes the management response for each evaluation recommendation addressed to the Commissioner and identifying follow-up actions.</li><li>■ Ensures that the output of the evaluation is utilized in program or project improvement.</li></ul>

Person/Unit	Roles and Responsibilities
Evaluation Manager	<ul style="list-style-type: none"> <li>■ Leads the development and finalization of the evaluation plan and ToR, particularly the methodology and approach with respect to the objectives and questions identified by the Evaluation Commissioner.</li> <li>■ Manages the procurement of the external evaluator/s in tandem with the agency's Bids and Awards Committee and Procurement Unit. <ul style="list-style-type: none"> <li>● Manages the contractual arrangements, the budget, and the staffing needs of in the evaluation.</li> <li>● Safeguards the independence of the exercise.</li> </ul> </li> <li>■ Convenes and provides coordination support to the Evaluation Reference Group.</li> <li>■ Briefs the evaluators on the ToR and the quality assurance requirements during the inception meeting.</li> <li>■ Lends support to the evaluators throughout data collection and report preparation.</li> <li>■ Ensures that the evaluators have access to the required data and key informants.</li> <li>■ Leads the review and quality assurance of key evaluation products by providing comments on the inception report and the draft report. Responsible for consolidating comments on these evaluation products from the ERG and the Evaluation Commissioner.</li> <li>■ Performs contract management responsibilities.</li> <li>■ Performs first-line quality assurance reviews for each evaluation product (see Section F).</li> </ul>

Person/Unit	Roles and Responsibilities
	<ul style="list-style-type: none"> <li>■ Follows through on the Management Response process by ensuring that all recommendations have a response from the appropriate agencies or units and that the same implement their committed actions.</li> <li>■ Ensures that evaluations are translated into communication products.</li> </ul>
Evaluation Reference Group	<ul style="list-style-type: none"> <li>■ Reviews and comments on draft TOR, specifically the scope, objectives, and key evaluation questions.</li> <li>■ Participates in the evaluation by providing data and acting as key informants.</li> <li>■ Reviews and comments on the inception report and the draft evaluation report.</li> <li>■ As appropriate, and if identified in the recommendations section, takes responsibility for preparing a management response to key issues and recommendations raised in evaluations, and identifying key follow-up actions, responsible units for implementation, and estimated completion dates for these actions.</li> </ul>
Evaluation Team	<ul style="list-style-type: none"> <li>■ Undertakes the evaluation in line with the public sector evaluation norms and standards and ethical guidelines, and according to the ToR. This includes: <ul style="list-style-type: none"> <li>● Developing an evaluation matrix as part of the inception report.</li> <li>● Drafting reports and briefing the Evaluation manager, evaluation commissioner and the evaluation reference group on the progress and key findings and recommendations.</li> </ul> </li> </ul>

Person/Unit	Roles and Responsibilities
Peer Reviewers	<ul style="list-style-type: none"><li>■ Provides an independent and anonymous review of the quality of the evaluation report particularly in terms of the quality of the application of methodology, correctness of the findings, usefulness of the recommendations, and other aspects.</li></ul>



**Evaluation Oversight:** Responsible for approving the agency’s evaluation agenda, including its evaluation budget, to ensure that the prioritized studies are relevant to the needs and priorities of the agency. The role of evaluation oversight may be assumed by the agency secretary, undersecretary, or a committee composed of members of the agency’s management committee.



**Evaluation Commissioner:** The lead official or unit responsible for the design and implementation of the program or project to be evaluated will take a supporting role in the implementation of an evaluation. The commissioning office typically is the program/project management unit (PMU) or the office to which the PMU reports.

To ensure independence and credibility of an evaluation they will not manage an evaluation directly but will support the identification of key evaluation questions and the development of the evaluation plan and TOR, support the conduct of the evaluation by facilitating the data collection mission, and ensure that the evaluation is utilized by developing and implementing management responses.



**Evaluation Manager:** The management of an evaluation should be separate from project management itself. Where an agency has an M&E or evaluation unit in place, then they should take the evaluation management role. Otherwise, this role is delegated to the M&E or evaluation focal/s of the planning bureau. While the evaluation manager will be responsible for final approval of all aspects of the evaluation process – from initiation to utilization, including acceptance of evaluators’ outputs – the evaluation commissioner retains a supporting role throughout the evaluation and, for the evaluation manager. Hence, under no circumstance should an evaluation manager be the same person or unit as the evaluation commissioner, and vice versa.



**Evaluation Reference Group or Evaluation Review Committee:** All evaluations should have a group of stakeholders that can support the evaluation process by giving expert advice and technical comments at key stages in the evaluation process. This ensures the transparency and credibility of the resulting evaluation. An evaluation reference group is composed of 5 to 9 members which include the evaluation commissioner, the evaluation manager, other relevant units of the agency, and other agencies and stakeholders as may be needed.



**Peer Reviewers:** All evaluations should be reviewed by an independent evaluation and/or subject-matter specialist to ascertain the quality of the evaluation design and outputs. The peer reviewer may be an individual or institution assigned to provide an independent and anonymous commentary and validation of the evaluation results. The peer reviewer is known only to the evaluation manager. The peer reviewer may be sourced through procurement (i.e., as a consultant) or by request (e.g., to a university/research organization or development partner).



- B. Once the evaluator (firm, individual consultant, or academic/research institution) has been procured, an inception meeting shall be held with at least the evaluator, the evaluation manager, and the evaluation commissioner to confirm the specifications and schedule under the TOR and the evaluator's bid proposal; and to discuss other arrangements such as access to data and key informants, mode of submission and communication, and other considerations.
- C. In cases when the agency opts for an internal evaluation, the staff performing the role of the evaluator may not concurrently perform the role of Evaluation Manager or Evaluation Commissioner. It will still be ideal for the internal evaluator to be governed by a formal TOR with specified outputs, quality standards, and timelines.
- D. Based on these discussions, the evaluator shall prepare and submit an Inception Report (see Template 6) that spells out the methodologies to be used to address each evaluation question and consistent with the TOR, the evaluator's implementation plan, and requirements from the evaluation manager, evaluation commissioner and other key stakeholders. The evaluation manager, in consultation with the evaluation commissioner and the ERG, shall review and accept the inception report. The inception report may be presented before the ERG in a meeting prior to revision and acceptance.

## 5.2 Managing risks

- A. Not all evaluation processes will run smoothly, no matter how meticulously planned. But evaluation managers can mitigate risks in the course of implementing the evaluations through the following steps:
  - Ensure a clear and regular communication among the Evaluation Manager, evaluation team, program manager, and key stakeholders in all stages.
  - Spot any conflicting stakeholder interests early on and find ways to minimize tensions between stakeholders when the evaluation findings are released.

- Ensure flexibility in program implementation, and the alignment of implementation with both program and evaluation objectives (Queensland Trade and Treasury, 2014).

- B. When the evaluation is not progressing as planned, the evaluation team should identify roadblocks, and to address them as soon as possible. This requires good project management, strategic implementation, and a willingness to call for a pause to determine what is wrong and to fix it.
- C. The evaluation manager should maintain a risks log and update it at key points of the evaluation (see Section 4.4). The evaluation manager should discuss risks and measures to mitigate risks or address risks that have been actualized with the evaluation commissioner and, if necessary, the evaluation oversight.

## 5.3 Ethical behavior

- A. Individuals engaged in designing, conducting, and managing evaluation activities shall abide by the Code of Conduct and Ethical Standards for Public Officials and Employees (Republic Act No. 6713) and the principles, norms, and standards spelled out in section 1.6 of these guidelines.
- B. Evaluators shall respect the implementing agencies' and individuals' right to provide information in confidence, and that ensure sensitive data they shared cannot be traced to them. (UNEG, 2016).

Evaluators must ensure that those involved in evaluations will have an opportunity to review and approve the statements attributed to them. Evaluators shall be sensitive to the cultural, social, and economic environment of all stakeholders, and conduct themselves in a manner that is fair and appropriate to this environment. Evaluators shall be accountable for their performance and their products.



C. To maximize objectivity and minimize the potential for bias, implementing agencies shall ensure the impartial conduct of evaluations. While their active involvement is typically necessary to effectively conduct an evaluation, managers and other stakeholders of a program under evaluation should not be allowed to influence evaluation findings. Good evaluation practice involves a constructive mix of internal and external stakeholders so that program knowledge can be brought to bear while potential conflicts of interest can be identified and prevented.

D. Additionally, those engaged in designing, conducting and managing evaluations should protect the privacy of individual subjects of evaluations and their sensitive information in line with the Data Privacy Act (R.A. 10173).

Evaluation managers and evaluators shall therefore establish research protocols for ensuring data privacy, such as data anonymization, adequate data security, non-disclosure and confidentiality agreements with stakeholders from the start of research, etc.

## 5.4 Evaluation reporting

A. To ensure transparency and accountability to all stakeholders, and inspire greater confidence in agencies and the government at large, all final evaluation reports (see Templates 7 and 8) shall be disseminated based on the parameters and conditions set in the Guidelines, and shall include the following:

- A description of the program being evaluated
- Adequate background and context, including the purpose of the evaluation and the evaluation issues and questions
- A description of the evaluation methodology, including limitations and the approaches adopted to mitigate limitations

- The identities of the evaluation team's principal members and, if applicable, the steering committee or other advisory bodies
- A concise statement of the evaluation findings, along with a description of the evidence on which each finding is based
- Conclusions derived from a synthesis of the findings
- A clear distinction between evaluator opinions (i.e., recommendations) and statements of established fact (i.e., findings)
- A management response from the agency head, describing actions that will be taken to respond to each recommendation.

B. The findings of an evaluation shall be used to make evidence-informed policy decisions, to exact accountability, and to improve existing programs. Decisions can be made at any time during the evaluation for continuous improvement of program implementation and management, or as guidance for designing future programs. (Queensland Trade and Treasury, 2014).

C. Different audiences will take evaluation findings differently. Hence, the approach for reporting on and disseminating evaluation findings should suit audience needs and expectations (See section 6 for additional guidance).

D. Regardless of the audience, good evaluation reports have the following characteristics (Western Australia Government, 2015):



### Results are reported fairly and accurately

- Make assumptions and value judgments explicit.



- Present data in a comprehensive way instead of selective.
- Avoid overgeneralizing results. Ensure the results indicate the stakeholders to whom the results apply, and the time frame for which the results will likely hold true.
- Avoid mistaking data correlation for causality<sup>32</sup> when there is not enough evidence to draw a given conclusion.

#### **The report is user-friendly**

- Be concise and use plain language in reporting evaluations. Avoid jargon.
- Provide background information to quantitative results to aid interpretation.
- Break up graphs and tables of numerical data with qualitative feedback that illustrates the points that the numbers tell.



#### **The report is produced in a timely manner**

- Strategize the release of information i.e., at the stages in planning, budgeting, and policymaking wherein findings can best influence decision-making.



#### **The report considers the ethical and political sensitivities and risks attached to the evaluation**

- Ensure truthfulness and accuracy while keeping in mind the values and sensitivities of both the community and the key stakeholders involved
- Provide clear guidance on the reliability and scope of results and the context under which they should be interpreted.

The evaluation manager, in consultation with the evaluation commissioner and the ERG, shall review and accept the evaluation outputs as spelled out in the TOR. The evaluation reports may be presented before the ERG in a meeting prior to revision and acceptance.



# UTILIZING EVALUATIONS

A. The value of an evaluation lies in how it can influence decisions and drive action. Useful evaluation reporting lays down options, identifies program strengths and weaknesses, and provides information on roadblocks to and opportunities for program improvement.

B. It is important to ensure that the findings are used to improve the delivery of interventions being evaluated. The following are some ways to promote the use and ownership of findings to inform decisions (Western Australia Government, 2015; Stetson, 2008; IDRC, 2003):

- Communicate the findings, particularly to decision-makers and key beneficiaries, throughout the evaluation process. Do not wait for the end when the evaluation report is produced.
- Align the release of findings with decision-making cycles.
- Link the findings to key outcomes and their indicators, as spelled out in development and investment plans.
- Present findings in an accessible, understandable, and culturally-sensitive manner.  
Use a variety of formats such as graphs and tables to illustrate key points and findings; and tailor them to each stakeholder group.
- Use the results to present an argument—including models or alternative proposals.

- Set up a process for discussing and designing an action plan in response to findings and recommendations. Create a tracker to follow up on compliance to and progress of actions taken.

## 6.1 Management response and implementation tracking

A. All evaluation recommendations require a corresponding Management Response (see Template 9), and for each accepted recommendation, a set of key follow-up actions illustrating how the recommendation will be addressed and when.

If the evaluation commissioner, manager, or members of the evaluation reference group disagrees with a recommendation, then they can state their respective justification in the management response. The agency's evaluation unit monitors the implementation of key actions and reports annually on the scale and pace of implementation.

B. Agencies or units identified in the recommendation section of the evaluation report are responsible for preparing a management response to each recommendation in the evaluation, including identifying key follow-up actions, responsible units for implementation of the key actions, and estimated completion dates.

These are uploaded to the online government evaluations portal within 3 months from the finalization of the evaluation study.

C. Management responses are important to ensure that evaluation findings are used and recommendations are acted upon. Lessons obtained from the evaluation and the management responses also serve as a rich resource for the design of new programs or projects.

D. Embedded in the Management Response Template is an implementation plan. The

implementation plan leads to enhancing the program or project by allowing the agencies to spell out their action points in response to recommendations and the timeline to carry out such actions. An implementation plan:

- Provides a process for thinking through critical components of a program.
- Provides clarity about possible steps to take in light of the findings and recommendations.
- Allows for anticipation of program challenges and enumerates the critical steps in advance.
- Provides a common understanding among stakeholders, particularly the agency and staff implementing the program—and their roles in implementing the actions.
- Clearly articulates the goal of the program (Western Australia Government, 2015; Urbis, 2011).

- E. The management responses and follow-through actions must be endorsed by the heads of the responsible units and of the agency concerned. This ensures commitment by the agency leadership to implement the recommendations.
- F. The Evaluation Task Force Secretariat is responsible for consolidating and reporting data on the status of management response implementation on an annual basis.

## 6.2 Communication and dissemination

- A. Communication plays a critical role in ensuring the success of an evaluation. While it is important to describe and explain the results of programs or projects as completed or implemented, the real impact is in the recommendations, particularly in how they can inform future programming decisions and drive results that

change lives. Hence, strategies for conveying evaluation information and messages should entail not only the communication of results, but communication for results

- B. All evaluation reports must be publicly disseminated following points d. and f. below. However, to protect its integrity, the evaluation may only be publicly discussed when ideally the evaluation report has been completed, or at least data collection has been completed and a draft report has been accepted. In the case of the latter, preliminary findings that have been defined with a degree of confidence may be communicated to a limited set of stakeholders for feedback and recommendations.

- C. Following are suggestions for presenting recommendations to key stakeholders (Western Australia Government, 2015; IDRC, 2003):

- Break down evaluation findings into three to five main points that are critical for each key stakeholder group. Provide recommendations based on these findings and develop a plan to ensure they are implemented and tracked for progress.
- Build ownership by involving key stakeholders in the development, finalization, and tracking of recommendations. Present recommendations in understandable formats and accessible channels.

Address them to appropriate stakeholders—i.e., those who will benefit from, and can contribute to making decisions and future actions out of these recommendations. Ensure they are context-appropriate, feasible, and practical.

- Link recommendations to the evidence

- D. To promote transparency and access to information, all evaluation reports and communication products shall be published, at the minimum, in the agency’s website and in the National Evaluation Portal (currently under development) that is managed by NEDA. Agencies are encouraged to disseminate their evaluation reports through various technical and popular means, including but not limited to media interviews and releases, social media, and stakeholders’ forums

- E. Agencies should likewise come up with a communications plan that spells out the communication formats and strategies to be used in reporting evaluation findings and recommendations, as well as the channels where they are to be disseminated, and the production timetable. It is imperative to consider all stakeholders in crafting messages and planning for formats and modes of communication.
- F. The publication of evaluation reports and the disclosure of related data and documents shall be compliant with the provisions under the Executive Order No. 2, s. 2016, on Freedom of Information. 4 The privacy of individual beneficiaries of programs that were covered by the evaluation shall be protected in line with RA 10173 or the Data Privacy Act of 2012. Where parts of an evaluation report fall under the exceptions of the FOI, the remainder of the report shall still be published. Where the agency deems that the whole report falls under the exceptions of the FOI, the agency shall report the said determination to NEDA.

## 6.3 Planning for the next evaluation

- A. Evaluation results shall inform how the program or project could be better monitored and evaluated. It is recommended that an evaluation should identify gaps in the collection of program or project data, ways to address these gaps and introduce innovations, and recommend the appropriate level of detail and frequency for administrative data collection without unduly adding to administrative burden. Improved monitoring and data collection systems improve the quality of future evaluations.

<sup>4</sup>Operationalizing in the Executive Branch the People's Constitutional Right to Information and the State of Policies to Full Public Disclosure and Transparency in the Public Service and Providing Guidelines Therefor



## QUALITY ASSURANCE

- A. The quality assessment of an evaluation report ensures the systematic review of the quality of an evaluation's design and its resulting findings, as well as the robustness of its conclusions and recommendations (UNDP, 2016).

The quality assurance criteria are applied to all evaluations commissioned by the national government. It serves as guidance for evaluation managers in the preparation of ToRs and in the review of the draft evaluation report before they are finalized.

- B. Mechanisms for quality assurance should be in place ideally throughout all phases of the evaluation cycle.

As such, and at the minimum, the ToR, draft report, and final report shall be subjected to a rigorous review. The key parameters of a quality assessment draw upon the basic requirements for acceptable outputs. Quality assessment considerations usually include the following (See Table 8):

Table 8. Quality assurance considerations

Source: UNDP (2019).

Evaluation component	Quality assurance considerations
Terms of Reference (ToR)	<p>Did the ToR appropriately and clearly outline the purpose, objectives, scope, and key questions for the evaluation?</p> <p>Have roles and responsibilities been adequately defined?</p>
Evaluation structure, methodology, and data sources:	<p>Is the evaluation structured well?</p> <p>Are the objectives, limitations, criteria and methodology fully described and appropriate?</p>
Findings, conclusions, and recommendations:	<p>Are findings appropriate and based on the evaluation criteria? Do they directly respond to evaluation questions?</p> <p>Do the conclusions draw from evidence presented in the evaluation and do they present logical judgments based on findings?</p> <p>Are the recommendations relevant to the subject and purposes of the evaluation, are they supported by evaluation evidence? Are the recommendations actionable?</p>

C. The evaluation manager carries out the first line of quality assurance, which entails checks to ensure key evaluation products – the ToR and the evaluation report – meet the quality criteria as set by governing evaluation norms and standards. As spelled out in Section E, the agency may tap the services of an independent peer reviewer. The ETF may also opt to subject evaluations to an independent quality assurance process on its own or through a third party.

## References

Asia-Pacific Community of Practice on Managing for Development Results. (2011). Framework for Results-based public sector management and country cases. Manila: Asian Development Bank. Department of Budget and Management. (2016). National Budget Memorandum 565: Adoption of a results-based monitoring, evaluation and reporting policy. Manila: DBM.

\_\_ (2017). National Budget Memorandum 565: Adoption of a Program Classification Expenditure-based Performance Informed Budgeting for the preparation of the Proposed National Budget for fiscal year 2018. Manila: DBM.

EvalPartners. (2016). Global evaluation agenda 2016-2020 (Retrieved 14 April 2019 from <https://bit.ly/2UAFgEc>).

International Development Research Centre. (2003). Designing and conducting health systems research projects - Volume II: Data analysis and report writing. The Netherlands: KIT Publishers.

Interpeace. (2010). Peacebuilding how? Criteria to assess and evaluate peacebuilding (Retrieved 14 April 2019, from <https://bit.ly/2VJVHug>).

National Economic and Development Authority. (2014). Ex-post evaluation manual. Manila: NEDA.

OECD-Development Assistance Committee. (1991). Principles for evaluation of development assistance, France: OECD Publications.

\_\_. (2002). Glossary of key terms in evaluation and results based management. France: OECD Publications.

\_\_. (2007). Encouraging effective evaluation of conflict prevention and peacebuilding activities. France: OECD Publications.



\_\_\_ (n.d.). DAC criteria for evaluating development assistance (Retrieved 25 February 2019, from <https://bit.ly/1TXgXX2>).

Preskill, H., & Jones, N. (2009). A Practical guide for engaging stakeholders in developing evaluation questions. New Jersey: Robert Wood Johnson Foundation Evaluation Series (Retrieved 27 February 2019, from <http://rwjf.ws/2x1sfaL>).

Republic of the Philippines. (2013). Republic Act 10625: Philippine statistical act of 2013. Manila, Philippines.

\_\_\_ (1989). Republic Act 6713: Code of conduct and ethical standards for public officials and employees. Manila, Philippines.

\_\_\_ (2016). Executive Order No. 2, s. 2016: Operationalizing in the Executive Branch the people's constitutional right to information and the State of policies to full public disclosure and transparency in the public service and providing guidelines therefor. Manila, Philippines.

Rosenstein, B. (2015). Mapping the status of national evaluation systems (2nd Ed.). DOI: 10.13140/RG.2.1.2008.0809.

Stetson, V. (2008). Communicating and reporting on an evaluation: Guidelines and tools. United States: Catholic Relief Services.

The State of Queensland (Queensland Treasury and Trade). (2014). Queensland Government program evaluation guidelines. Brisbane: Queensland Treasury and Trade, Economics Division.

United Nations Development Programme. (2016). Quality assessment system for decentralized evaluation reports. Retrieved 25 February 2019 from <http://web.undp.org/evaluation/documents/guidance/UNDP-Quality-Assessment-System-for-Decentralized-Evaluation.pdf>

\_\_\_ (2018). Evaluation guidelines. New York: UNDP Independent Evaluation Office.

UNDP Internal Evaluation Office. (2015). Bangkok principles on national evaluation capacity for the Sustainable Development Goals era. (Retrieved 14 April 2019, from <https://bit.ly/1QHMBDn>.)

United Nations Evaluation Group. (2013). Integrating human rights and gender equality In evaluations. New York: United Nations Evaluation Group.

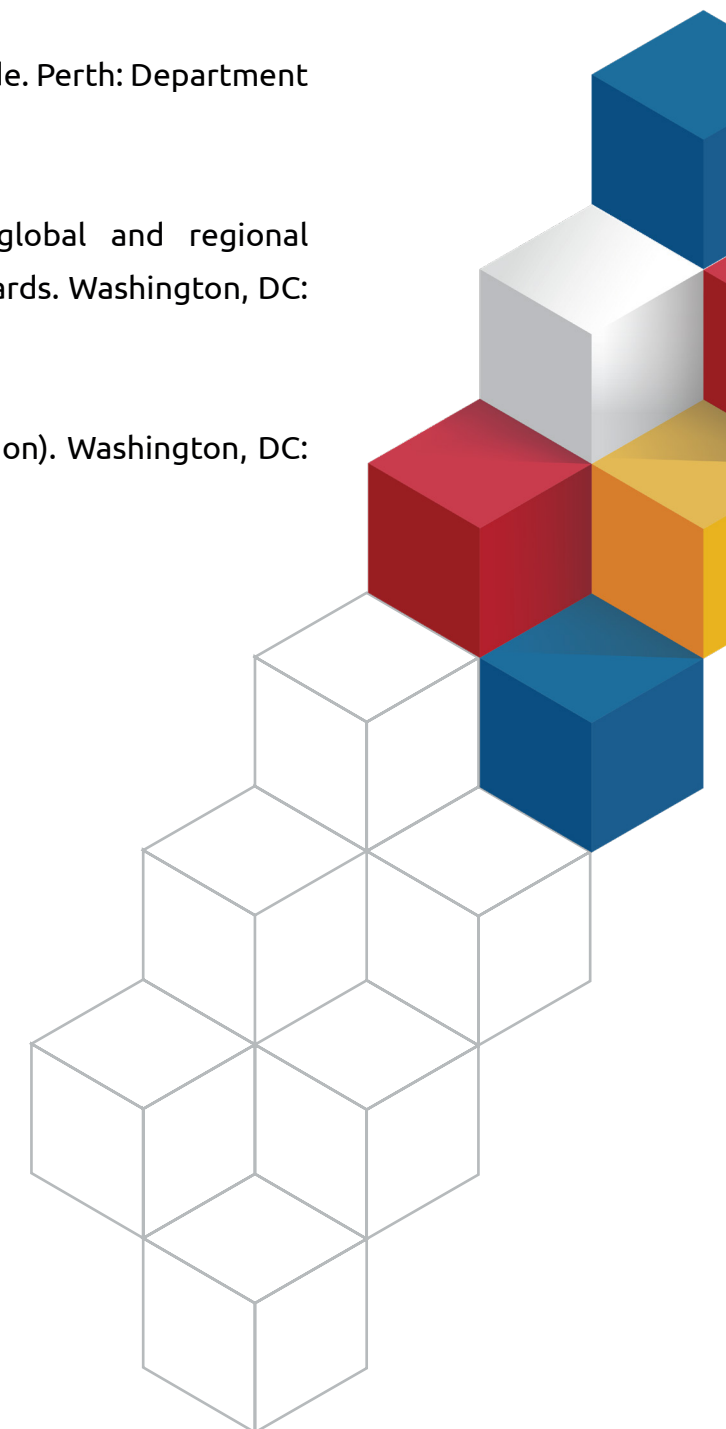
\_\_\_ (2019). Norms and standards for evaluation. New York: United Nations Evaluation Group.

Urbis Social Planning and Research. (2011). Implementation plan for evaluation of Keep Them Safe for the NSW Department of Premier and Cabinet. Australia: Urbis.

Western Australia Government. (2015). Evaluation guide. Perth: Department of Treasury, Program Evaluation Unit.

World Bank. (2007). Sourcebook for evaluating global and regional partnership programs: Indicative principles and standards. Washington, DC: Inter-American Development Bank and World Bank.

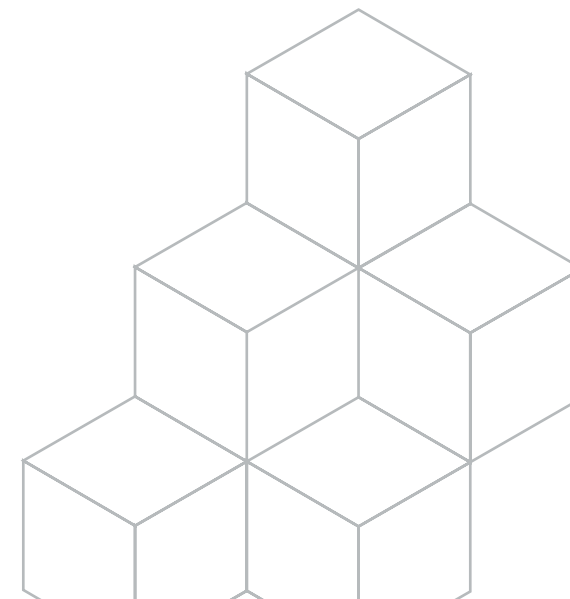
\_\_\_ (2016). Impact evaluation in practice (Second Edition). Washington, DC: Inter-American Development Bank and World Bank.



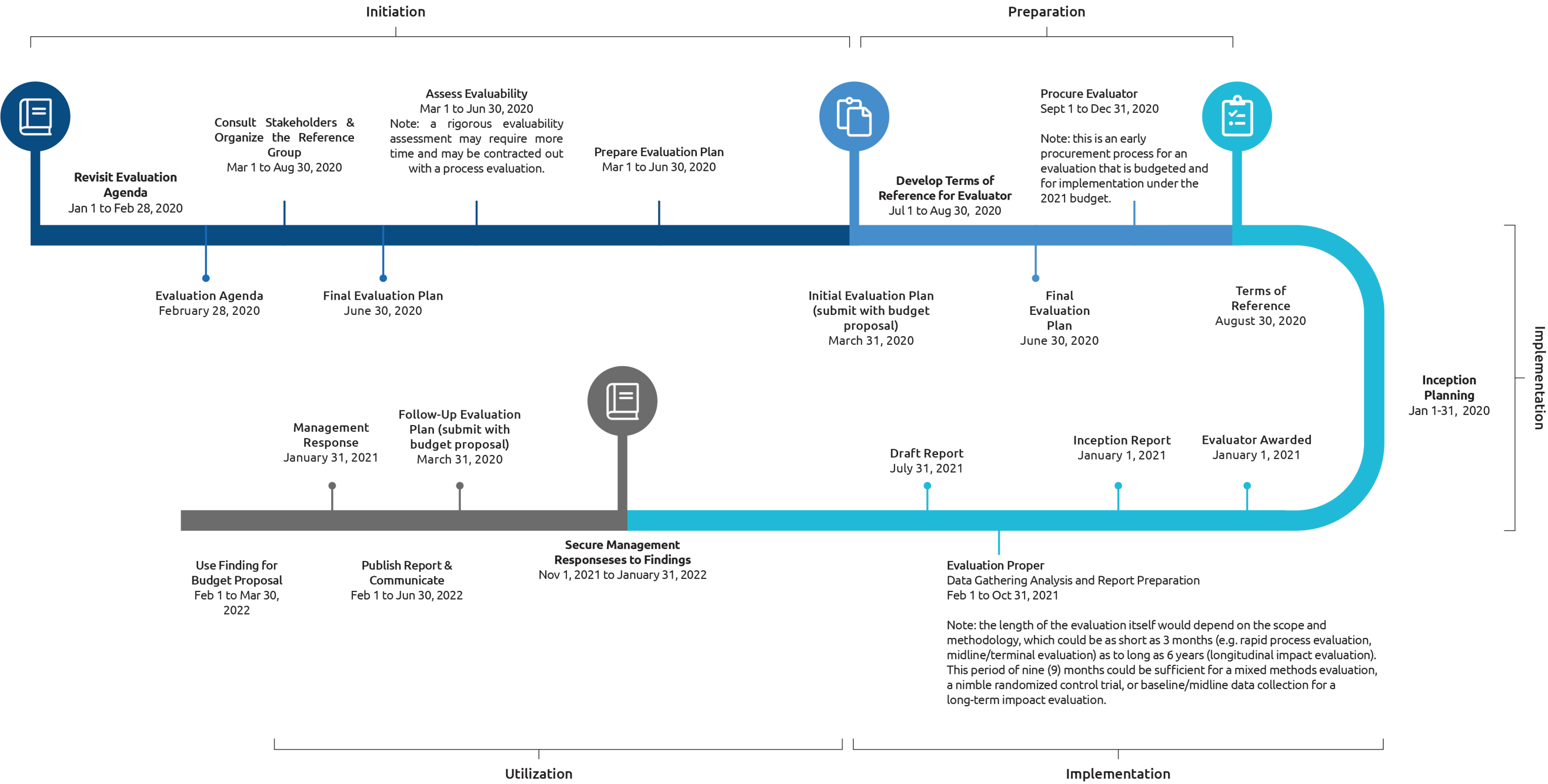


# Annexes

- Illustrative Timeline for Evaluations
- Evaluation Approaches and Methods
- Data Collection
- Sample Evaluation Questions
- Guide to Procuring Evaluations



# Annex A: Illustrative timeline for an evaluation



# Annex B: Evaluation approaches and methods<sup>5</sup>

In broad terms, evaluation approaches can be grouped into four following broad categories:



Different methods may be used to collect and analyze data and inform findings for any of these evaluation approaches. The ideal choice will depend on available data, timing and resourcing constraints, and how robust the findings need to be to inform decision-making.

It is highly unlikely that one evaluation method or technique can isolate the impacts of a program from all other effects, or provide a complete picture of program impacts. Where possible, evidence collected from a range of evaluations should be pulled together into a body of evidence to form a comprehensive view of program impacts.

The sections below outline some of the main methods that can be used for program evaluation and discuss some of the key things to consider when selecting a preferred approach.

## B.1 Evaluating design, implementation and delivery

A process evaluation focuses on how a program is being or was implemented and whether it is providing services as intended (focusing on output delivery rather than on outcomes).

A process evaluation attempts to answer the following types of questions:

- Is the program being implemented as intended?
- Are the activities being delivered as intended?
- Are participants or beneficiaries being reached as intended?

While several factors can inform whether a program has been implemented successfully, a successful program generally have:

- Been implemented on schedule and on budget
- Delivered the agreed outputs, at an acceptable quality, to the agreed audience
- Met the requirements of stakeholders.

Information obtained through a process evaluation can shed light not only on what is happening, but also why. Process evaluations can assist in identifying intended or unintended implementation characteristics that may have positively or negatively affected program delivery and impacted on program outcomes.

A process evaluation assesses issues such as the adequacy of administrative process, management practices and service delivery mechanisms, and the acceptability of services to program stakeholders.

A process evaluation can occur at any stage of program delivery and potentially on multiple occasions, particularly for ongoing programs, and can include both qualitative and quantitative methods. Appropriate methods will depend on the purpose of the evaluation and the level of evidence required (See Table B.1).

<sup>10</sup>This annex is based on Attachment A in Queensland Government (2014) and Section 2 in UNDP.

## B. 2 Evaluating progress on outcomes

Outcome evaluations capture contributions towards the achievement of outcomes as identified in the country’s development plan. Outcome evaluations are undertaken to:

- Provide evidence to support accountability of agencies and coordinating bodies
- Provide evidence of contribution to outcomes
- Guide performance improvement by identifying current areas of strengths, weaknesses, and gaps, especially in regard to:
  - Appropriateness of coordination strategy
  - Impediments to the outcome being achieved
  - Mid-course adjustments (for Outcome Midterm Reviews)
  - Lessons learned for the next planning/programming cycle

Table B.2: Differences between project and outcome evaluations

Evaluation questions	Description	Methods and data sources
Focus	■ Generally speaking, inputs, activities, and outputs (if and how project outputs were delivered within a sector or geographic area and if direct results occurred and can be attributed to the project).	■ Outcomes (whether, why, and how the outcome has been achieved; and the contribution of the agency/ies to change in a given development situation)
Scope	■ Specific to project objectives, inputs, outputs and activities.  Also considers relevance and continued linkage with outcome.	■ Broad, encompassing outcomes and the extent to which programs, project, soft assistance, partners’ initiatives and synergies among agencies and partners contributed to its achievement.
Purpose	■ Project based, to improve implementation, to re-direct future projects in the same area or to allow for upscaling of project.	■ To enhance development effectiveness, to assist decision-making, to assist policy making, to re-direct future planning, to systematize innovative approaches to sustainable human development.

Adopted from UNDP (2018)

Table B.1: Key evaluation questions by type of evaluation

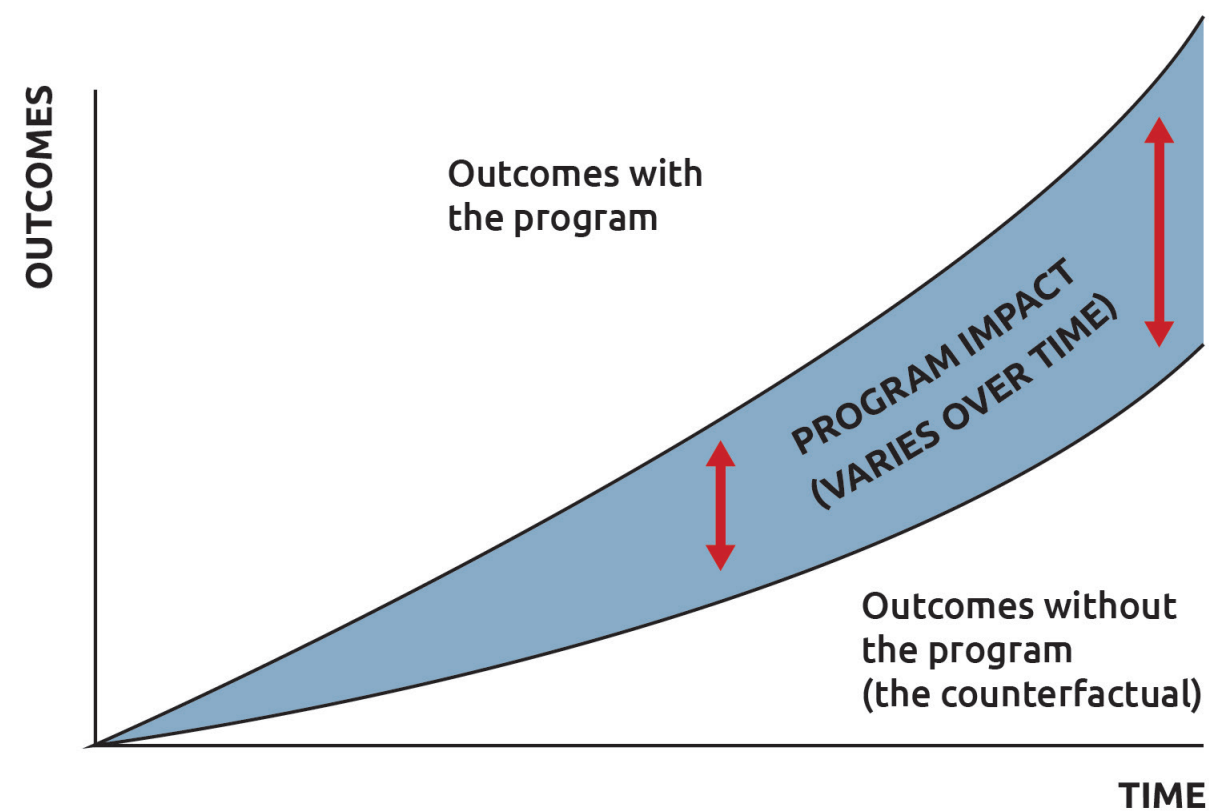
Evaluation questions	Description	Methods and data sources
How well has the program been established?	<ul style="list-style-type: none"><li>■ Description of program development compared with client needs, timeframes etc.</li><li>■ Quality of governance, relationships.</li><li>■ Influence of different factors and contexts.</li><li>■ Initial evidence of uptake.</li></ul>	<ul style="list-style-type: none"><li>■ Program reports, key informant interviews, consultations with managers or service providers, program forums.</li></ul>
How is the program being implemented?	<ul style="list-style-type: none"><li>■ Description of implementation processes by different providers and in different circumstances.</li><li>■ The extent that implementation processes met milestones and targets for outputs, timeliness, cost, participation and immediate outcomes.</li><li>■ The quality of outputs and immediate outcomes measured against standards and targets.</li><li>■ The pattern of outputs, uptake and immediate outcomes, by different sub-groups or in different contexts.</li><li>■ Client or customer satisfaction.</li></ul>	<ul style="list-style-type: none"><li>■ Program monitoring data and other program records.</li><li>■ Observation including photography and video.</li><li>■ Interviews, surveys or focus groups with managers, staff, program clients, referring agencies.</li><li>■ Consultations with managers or service providers.</li></ul>
Is the program being implemented well?	<ul style="list-style-type: none"><li>■ As above plus information about good practice in implementation processes</li></ul>	<ul style="list-style-type: none"><li>■ Expert review of program documents, or observations during site visits.</li></ul>

Source: NSW. 2014. Policy Makers Toolkit. Department of Premier and Cabinet.

# B.3 Evaluating effectiveness

Assessing a program’s effectiveness requires consideration of what would have happened in the absence of the program (the counterfactual) or what has happened after program implementation in comparison to what has occurred prior to implementation.

Figure A.1: Defining program impact



Most observed outcomes are affected by a range of factors, not just the program in question. As such, a rigorous evaluation of program effectiveness will make a concerted attempt to isolate the effect of the program from all other potential influences<sup>6</sup>. For example, an evaluation of a program aimed at getting the unemployed back to work would need to consider if the participants found employment a result of the program or from other factors, such as changing labor market conditions.

An impact or outcome evaluation focuses on the outcomes of a program and attempts to ascertain whether a program was successful in achieving its intended impact (focusing on outcomes rather than outputs). An impact evaluation should ask:

- Did the program achieve its stated objectives?
- How much of the changes could be attributed to the program as opposed to other factors?
- What was the magnitude of the impact achieved by the program?
- Were there any unintended consequences, and if so, how significant were they?

Impact evaluations may be carried out using a variety of methods, which can be broadly classified into three categories, depending on the level of control over variables and the ability to determine a causal relationship between a program’s activities and observed outcomes:



## Experimental design

Based around the selection of a comparison and treatment group, through random assignment, prior to the implementation of a program. The treatment group is then exposed to the program, while the comparison group is not. A randomized-control trial is an example of an experimental design. Experimental design requires a very high level of control which can be very difficult to achieve outside clinical studies.



## Quasi-experimental design

Consists of the evaluator selecting a comparison group, that is matched with the treatment group on certain characteristics (either before or after a program has been implemented) to provide a comparator for the treatment group. Quasi-experimental design offers less control than experimental design, in that subjects are selected from preexisting groups; e.g. criminal offenders or people with a disability. Difference-in-difference, regression discontinuity and matching samples are examples of quasi-experimental designs.

<sup>11</sup> Government Social Research Unit. 2011. The Magenta Book: Guidance for evaluation. London: HM Treasury available online at <http://bit.ly/1zABTG7>.



The most appropriate method will be determined by the resources and time available for evaluation, the feasibility of implementation and the strength of evidence required from the evaluation findings (Table B.3).

Table B.3: Summary of impact evaluation methods

Evaluation design	Description	Methods and data sources
<b>Experimental</b> <i>(such as a randomized control trial)</i>	<ul style="list-style-type: none"><li>■ Involves gathering a set of equally eligible individuals willing to participate in the program, randomly assigning them to either a treatment group (those affected by the program) or a control group (unaffected by the program).</li></ul>	<ul style="list-style-type: none"><li>■ Attempts to carefully control confounding factors in the design and hence improve the ability to attribute changes to the intervention itself.</li><li>■ Can be resource intensive and difficult (or unethical) to provide an intervention to one group and not another.</li><li>■ Can also be very difficult to achieve the required level of control outside of clinical settings.</li></ul>
<b>Quasi-experimental</b> <i>(such as difference-indifference, regression discontinuity design, and matching)</i>	<ul style="list-style-type: none"><li>■ Consists of constructing a comparison group using matching or reflexive comparisons either before or after program implementation (prospective or retrospective evaluation).</li></ul>	<ul style="list-style-type: none"><li>■ Can allow evaluators to draw on existing data sources, making this method quicker and cheaper to implement.</li><li>■ The matching methods can be statistically complex, and the methodology may not completely solve the problem of selection bias, which can reduce the reliability of the results.</li><li>■ Makes some attempt at controlling confounding factors in the design to improve the ability to isolate intervention effects.</li></ul>

<b>Non-experimental</b> <i>(methods that rely on interpretation, observation or interactions such as realist, theory-driven and participatory approaches)</i>	<ul style="list-style-type: none"><li>■ Includes a very wide range of different types of evaluations and other studies that do not compare results for the same group over time or different groups of participants.</li></ul>	<ul style="list-style-type: none"><li>■ The evaluation design can be relatively cheap and quick to implement.</li><li>■ A number of difficulties in ensuring the validity of findings mean this design tends to provide weak evidence about the effectiveness of a program.</li></ul>
--	--	---



**Non-experimental design**

Will not demonstrate a cause-and-effect relationship, but instead seek to understand how and why a program is effective (not just whether it is effective). Non-experimental designs include a very wide range of methods, including realist, theory-driven and participatory approaches, which rely on interpretation, observation or interactions to come to a conclusion. The most appropriate method will be determined by the resources and time available for evaluation, the feasibility of implementation and the strength of evidence required from the evaluation findings (Table B.3).

Considerations when selecting an appropriate method for analyzing impact include:

**Determining causality** - There are many factors that can compromise the ability of an evaluation to demonstrate that a relationship exists between a program’s actions and the observed outcomes. These factors are known as internal threats to validity, and include history, maturation and design contamination. Implemented appropriately, experimental and quasi-experimental designs, measuring prior to and throughout program implementation, will provide the best means of controlling for these factors.

- **Generalizability of results** - If the evaluation results are going to be used to inform decision around the likely effect of the program on different individuals, contexts or outcomes, the evaluation will need to control for factors that might reduce the generalizability of the evaluation findings. These factors are also known as external threats to validity.

- **Sample size** - In general, the larger the sample size, the easier it will be to separate out the impacts of a program from other random variations. It may be easier, or more cost effective, to have larger sample sizes for some methods over others.

- **Cost** - Different evaluation methods will require different time, resourcing and levels of expertise. Evaluations with controlled or quasi-experimental design will almost always cost more than non-experimental evaluations.

However, the additional cost can sometimes be quite low depending on the type and availability of data to be collected. Moreover, findings from a more extensive evaluation may lead to future cost-savings, through improved outcomes and more efficient use of resources. Nevertheless, program managers must anticipate these additional costs when budgeting for program implementation<sup>7</sup>. This reinforces the need for the evaluation plan to be completed during the design phase of the program, and for the program budget to separately identify evaluation costs.

- **Type of information provided** - Experimental and quasi-experimental designs explicitly seek to demonstrate cause and effect relationships (i.e., evidence of whether a program is effective in delivering outcomes).

While non-experimental or observational designs are less likely to be able to demonstrate the causal link between program actions and outcomes, they can often add value to an evaluation through their ability to inform an understanding of why events or outcomes unfolded in a particular way.

- **Complexity of the program** - For programs with multiple delivery strategies, the evaluation may need to examine whether individual elements of a program are effective, and how they contribute to the effectiveness of the program as a whole. A realist-driven approach to program evaluation may be required in such circumstances.

Realist evaluation seeks to understand which initiatives appear to be effective, for whom and under what circumstances, recognizing that there can be multiple pathways to any given outcome. Such an approach seeks to link the mechanisms triggered by a context to produce a particular outcome.

## B.4 Evaluating efficiency

A reliable impact evaluation focuses on demonstrating and quantifying the outcomes generated by a program, but it does not provide information on a program's efficiency and value for money.

The concept of efficiency has a number of dimensions: technical, allocative and dynamic efficiency. For the purposes of program evaluation:

- Technical efficiency requires that the program is delivered at the lowest possible cost
- Allocative efficiency requires that the program is delivered to the areas of greatest need, and provides the types of outputs or outcomes that recipients value most (for the given set of resources available)
- Dynamic efficiency means that, over time, the program continues to improve, by finding better or lower cost ways to deliver outcomes.

The extent to which it will be possible to examine all three types of efficiency in a quantitative framework will depend on the scope of the evaluation and the type of program being delivered. Where quantitative assessment of a particular type of efficiency is not possible, the efficiency aspects can be addressed through a qualitative discussion.



### Program-specific considerations

When designing a program, the program manager should consider whether there are ways to build in evaluation of a program's efficiency in delivering outcomes.

At a basic level, examination of a program's efficiency should consider whether:

- Changes to processes could result in more efficient service delivery
- There are elements of the program that are currently delivered by government that could be delivered more effectively by the private sector
- There are different solutions (such as alternate technologies) that could deliver the same outcomes more efficiently or effectively, or improved outcomes for the same cost

<sup>8</sup>Department for International Development, 2014, Assessing the Strength of Evidence, available online

- There are similar programs being delivered by other agencies or bodies, with potential to more strategically target efforts or pool resources
- the program is being delivered to the areas of greatest need
- the program is delivering the types of outputs desired by program recipients.

An examination of efficiency should also involve consideration of whether there are any factors impeding the achievement of more efficient service delivery and investigating ways these could be overcome (to ensure the program can remain efficient over time).

If a more cost-effective solution is identified and approved for implementation, the program manager would need to ensure that the implementation of the alternate solution was successful and that the effectiveness of the program in achieving outcomes was not unduly compromised (which may instigate the need for a revised evaluation plan).



**Whole-of-economy impact**

For programs that are large, complex, expensive, associated with a high degree of risk or with the potential for large behavioral impacts, it may also be necessary to consider whether the benefits of the program outweigh the costs.

*Economic evaluation* takes a wider scope than the program-specific considerations discussed above, aiming to address the broader question of whether the program is the most appropriate way of addressing the problem that triggered the introduction of the program. Economic evaluation considers the financial, environmental and social costs and benefits generated by the program against other delivery options to determine if the program provides the best solution from a whole-of-society perspective.

For example, a program manager who has been asked to deliver a treatment program to address a particular health issue could evaluate whether the program is cost-efficient and effective, while government more broadly should evaluate whether a treatment program is the best option, or if a preventative program may provide greater benefit or value for money. An economic evaluation can be used to inform decisions on whether an existing program should be continued or ceased and/or if the current mix of programs is efficient and provides value for money.

An economic evaluation involves systematically identifying, valuing, and comparing the costs and consequences of two or more alternative programs or interventions in a unified framework. Where appropriate, it is desirable to quantify the economic impacts of the program to the extent possible, while this information can then be combined with consideration of broader issues such as affordability, equity, acceptability, feasibility and the strength of the evidence base to determine whether a program should be implemented, continued, or ceased.

There are three main types of economic evaluation that assess the economic efficiency of programs: cost-benefit analysis, cost-effectiveness analysis and multi-criteria analysis (Table B.4).

Table B.4: Overview of economic evaluation methods

Evaluation design	Design features	Method assessment
Cost-benefit analysis	<ul style="list-style-type: none"><li>■ Compares the financial, economic, environmental and social costs of a program from the perspective of society as a whole.</li><li>■ Involves the comparison of one or more program scenarios with an alternative scenario based on estimates of what would have happened in the absence of the program.</li><li>■ Requires the monetization of both costs and benefits to the extent possible.</li></ul>	<ul style="list-style-type: none"><li>■ The most commonly accepted method for assessing the economic efficiency of a program.</li><li>■ Quantification of the full range of costs and benefits can be resource intensive.</li><li>■ Can ignore equity concerns, although such measures can also be assessed or discussed in qualitative terms as part of a cost-benefit analysis.</li></ul>
Cost-effectiveness analysis	<ul style="list-style-type: none"><li>■ Takes a similar approach to cost-benefit analysis but compares the costs of an intervention to a common physical unit of outcome (such as “number of lives saved”).</li><li>■ Requires the same outcome measure to be relevant for all programs being compared.</li></ul>	<ul style="list-style-type: none"><li>■ Can only be used in situations where programs have shared goals; when the primary objective of the evaluation is to identify the most cost-effective strategy from a group of alternatives that can effectively meet a common goal.</li></ul>

Evaluation design	Design features	Method assessment
Multi-criteria analysis	<ul style="list-style-type: none"><li>■ Includes a broad range of techniques that can be used to assess alternative options according to a variety of criteria that have different units (e.g. dollars, tonnes, and kilometers).</li><li>■ Involves assigning weights and scores to each criterion, to indicate the relative importance.</li></ul>	<ul style="list-style-type: none"><li>■ The method allows for both qualitative and quantitative criteria to be considered, however the subjectivity of weighting and scoring means there is a high likelihood of bias.</li></ul>

Regardless of the type of method selected, it is expected that an economic evaluation will include the following features:

- A broad view of costs and benefits, including environmental and social impacts, and longer-term effects
- Identification of all costs and benefits, even those that cannot be monetized, including consideration of both indirect and non-market costs and benefits
- Assessment of the impacts on different stakeholder groups (such as those directly impacted by the program, the government, and the wider community)
- Sensitivity analysis: information on all costs and benefits is rarely known with certainty, necessitating the need to test the sensitivity of the results to changes in key parameters
- The findings of an economic evaluation will generally be heavily dependent on the assumptions chosen to underpin the analysis. As such, any evaluation report should clearly set out the assumptions chosen and explore the impacts that the assumptions have on the findings (the results of sensitivity analysis). Ensuring that all assumptions made during the evaluation process are transparent is essential for the explanation and use of the evaluation findings.

# Annex C: Data Collection

A data collection plan should be embedded into program design and implementation during the program’s development. No single method of data collection is likely to be suitable to answer all evaluation questions. For most evaluations, a combination of both qualitative and quantitative information is used to evaluate a program. Choosing the right data to collect is key to getting valid information that stakeholders will perceive as useful for decision making.

To ensure reliability and consistency of data collected across different time points, all individual sets of data should be collected in the same way, using the same set of instruments (for example, a questionnaire) at each time point. Where one on one interviews, focus groups or broader surveys are used, an effort should be made to maintain consistency among respondents (for example, using the same people or target group). All interviewees and focus group participants should be assured that all contributions will be treated confidentially and will not be attributed to them.

The strength of evaluation findings is usually found in the bringing together of data from different sources. Data collection methods should be designed and scaled to the evaluation in accordance with the program’s size, risk and significance.

Collection of new, primary data for an evaluation may not always be practical. The use of existing or “secondary” data may provide an alternative means of answering the evaluation questions. Sources of secondary data include other studies previously conducted, administrative records as well as (though not limited to) official statistical sources. After identifying a secondary data source, it is important that its relevance and quality is determined by the program evaluation working group before using it as part of an evaluation.

The establishment and collection of baseline data before program implementation is recommended to enable meaningful comparisons where possible. If baseline data cannot be obtained, the use of benchmarking against similar programs or best practice research can be used as a baseline guide for the program evaluation.

The Philippines has subscribed to the 2030 Agenda for Sustainable Development, which elevates as a core principle the objective to “leave no one behind” (LNOB). This principle will require the use of disaggregated data to assess trends across different groups. Considerations when choosing the disaggregation of variables, include:

- Sex
- Age
- Location
- Migratory status
- Ethnicity
- Wealth Quintile (income or wealth index)
- Marital status

Methodological approaches to data collection may include some or all of the following (UNDP, 2019):



Document review of all relevant documentation this would include a review of:

- Program or Project Document
- Work and Financial Plans
- Activity Designs
- Quarterly and Annual Reports, such as the Budget and Financial Accountability Reports as required by COA-DBM Joint Circular No. 2014-1 and subsequent issuances
- Highlights of Project Board Meetings or their equivalent
- Other relevant documents



Semi-structured interviews with key stakeholders including lead government agencies or units, other involved agencies or units, development partners, civil society organizations, academe, representatives of beneficiary groups, others.

- Development of evaluation questions around relevance, coherence, effectiveness, efficiency and sustainability and designed for different stakeholders to be interviewed;
- Key informant and focus group discussions with beneficiaries and stakeholders; and
- All interviews should be undertaken in full confidence and anonymity. The final evaluation report should not assign specific comments to individuals.



Surveys and questionnaires, including participants in programs, other stakeholders at strategic and programmatic level.



Other methods such as outcome mapping, observational visits, etc.



Field visits and on-site validation of key tangible outputs and interventions.



Data review and analysis of monitoring and other data sources.

- Ensure maximum validity, reliability of data (quality) and promote use, the evaluation team will ensure triangulation of the various data sources.

The final methodological approach, including interview schedule, field visits, and data for use in the evaluation should be clearly outlined in the inception report of the evaluation and be fully discussed and agreed between the evaluation commissioners and managers, stakeholders and the evaluators.

# Annex D: Sample Evaluation Questions

The following sample evaluation questions are adopted from the NEDA-DBM Joint Memorandum Circular No. 2015-01 (National Evaluation Policy Framework Guidelines), as prescribed by evaluation organizations (OECD-DAC and UNEG) and adopted by development institutions (e.g., ADB, WB, and JICA). These questions may serve as a guide when formulating your list of evaluation questions.





## Relevance

### Responsiveness to stakeholder needs

1. To what extent does the program address the urgent needs of the stakeholders?
2. Did the project design and choice of outputs and activities properly reflect and address the needs of communities?

### Programmatic alternatives

3. Are there better ways to achieve the outcomes / impacts of programs / projects, or for programs / projects to contribute to related national priorities?

### Project design

4. How valid is the theory of change?
5. Were the planned and actual outputs and activities of the project consistent with the intended outcomes?



## Coherence

### Compatibility with other interventions, Internal or External.

6. Do other interventions and policies support or undermine the program/project being evaluated, and vice versa?
7. Do other interventions and policies support or undermine the program/project being evaluated?
8. Does the intervention add value without duplicating effort?

### Internal coherence

9. To what extent is the program/project harmonized with the other interventions of the government?

10. To what extent do program outcomes/impacts align with the achievement of national priorities and existing laws, including PDP sector outcomes?
11. How consistent is the intervention with the relevant international norms and standards adhered to by the government?

### External coherence

12. How consistent is the program/project with the interventions of other actors in the same context?
13. How relevant is the program/project international norms and standards adhered to by the government?
14. How does it complement, harmonize, and coordinate with such other actors' interventions?
15. Does the intervention add value without duplicating effort?



## Effectiveness

### Achievement of objectives

16. What intended outputs and outcomes / impacts (short-term outcomes, medium-term outcomes, and long-term impacts) were found, and to what extent can they be attributed to program activities?
17. How likely will the project / program contribute to the planned outcomes?
18. To what extent was the project implementation strategy appropriate to achieve the objectives?
19. What challenges have been faced? What has been done to address the potential challenges/problems? What has been done to mitigate risks?

### Unintended results

20. What unintended outputs and outcomes / impacts were found, and to what extent can they be attributed to program activities?

<sup>14</sup> Sample evaluation report with evaluation questions: <https://erc.undp.org/evaluation/plans/detail/1382>



### Coverage

21. Is the project reaching the intended beneficiaries, rights holders, and duty bearers?

### Value added

22. What value has the implementing agency or the project/program added?



## Efficiency

### Efficient delivery of outputs

23. Were the activities cost-efficient?
24. Was the utilization of resources optimized in terms of the realization of the program objective?
25. To what extent was resource utilization minimized in relation to the delivery of outputs?
26. Are the variances between planned and actual expenditures justified?

### Operational alternatives

27. Are there better, more efficient ways to deliver program outputs?

### Timeliness

28. Were the objectives achieved on time?



## Sustainability

### Continuity of benefits

29. How likely will the benefits of the project continue in the future?
30. To what extent did the benefits of a program continue after funding ceased?
31. Does the project have a continuity or exit strategy?

### Success factors

32. What were the major factors which influenced the achievement or non-achievement or sustainability of the program?



## Impact

### Overall effects of intervention vis-à-vis counterfactual

33. What changes or effects has the program / project brought?
34. What would have happened if the intervention had not taken place?

### Tangible effects vis-à-vis baselines

35. What real difference has the intervention made to the intended beneficiaries?
36. How do the changes fare from the initial state/ situation of the beneficiaries?

### Effects to stakeholders

37. How many people have been directly and indirectly affected by the intervention?



## Other criteria

### Gender Mainstreaming

38. To what extent have gender equality and the empowerment of women been addressed in the design, implementation, and monitoring of the project?
39. Does this answer to the Harmonized Gender and Development Agenda and/or the Gender Mainstreaming Framework?

## Responsiveness

40. Is the project's gender responsiveness classification for the project design and project implementation, management, monitoring, and evaluation representative of reality?

## Results

41. To what extent has the project promoted positive changes in gender equality and the empowerment of women?
42. Were there any unintended effects? (UNDP, 2019)



## Human Rights

### Participation and inclusion

43. How has the intervention helped in creating the conditions for active, free, and meaningful participation in the context of bringing development for all?
44. Which groups were able to participate in or benefit from it, and which ones were left out?
45. Were measures purposely put in place to support the participation of women and the marginalized?

### Equality and non-discrimination

46. Were all stakeholders given access to the intervention and the processes to avail it?
47. Were they treated fairly, and all their needs addressed regardless of gender, origin, age, or disability?

### Social transformation

48. Did the intervention allow for more dignified, humane, and equitable conditions among women and the marginalized? (UNEG, 2013)



## Peace

### Linkages

49. Do nationwide initiatives to bring peace and development seek to address the geographical (regional/international) and socio-cultural dimensions of the conflict?
50. Do the interventions involving leaders and decision-makers link effectively with efforts to engage on the grassroots level?

### Coverage

51. Do the interventions and resources adequately extend to all areas where conflicts happen?
52. Has sufficient attention been paid to emerging threats to peace?

### Consistency with conflict prevention and peacebuilding values

53. Are conflict- and culture-sensitive approaches practiced by implementation and engagement staff?
54. Are the interventions consistent with peacebuilding ends?
55. Are they inclusive enough to foster unity and dialogue among cultures, religions, and ethnic groups? (Interpeace, 2010; OECD-DAC, 2007).

# Annex E: Procuring Evaluation

All services required for evaluations, including but not limited to consultancy firms and individual consultants, shall be procured in line with the Government Procurement Reform Act (GPRA, R.A. No. 9184) and its latest Implementing Rules and Regulations (2016 IRR), or other applicable policies, such as international agreements.

As a policy, transparent and competitive procurement shall be undertaken with due regard to ensuring streamlined processes, accountability, and public participation.

Evaluations entail the procurement of consultancy services, specifically of management and related services. Whether firms or individuals are sought, consulting services give emphasis to the technical soundness of proposals compared to price.

Evaluations fall under the Management and Related Services category of consulting services. Competitive bidding shall be the default mode of procurement (GPRA IRR Sec 10), unless the use of alternative modes is deemed more advantageous to the government.

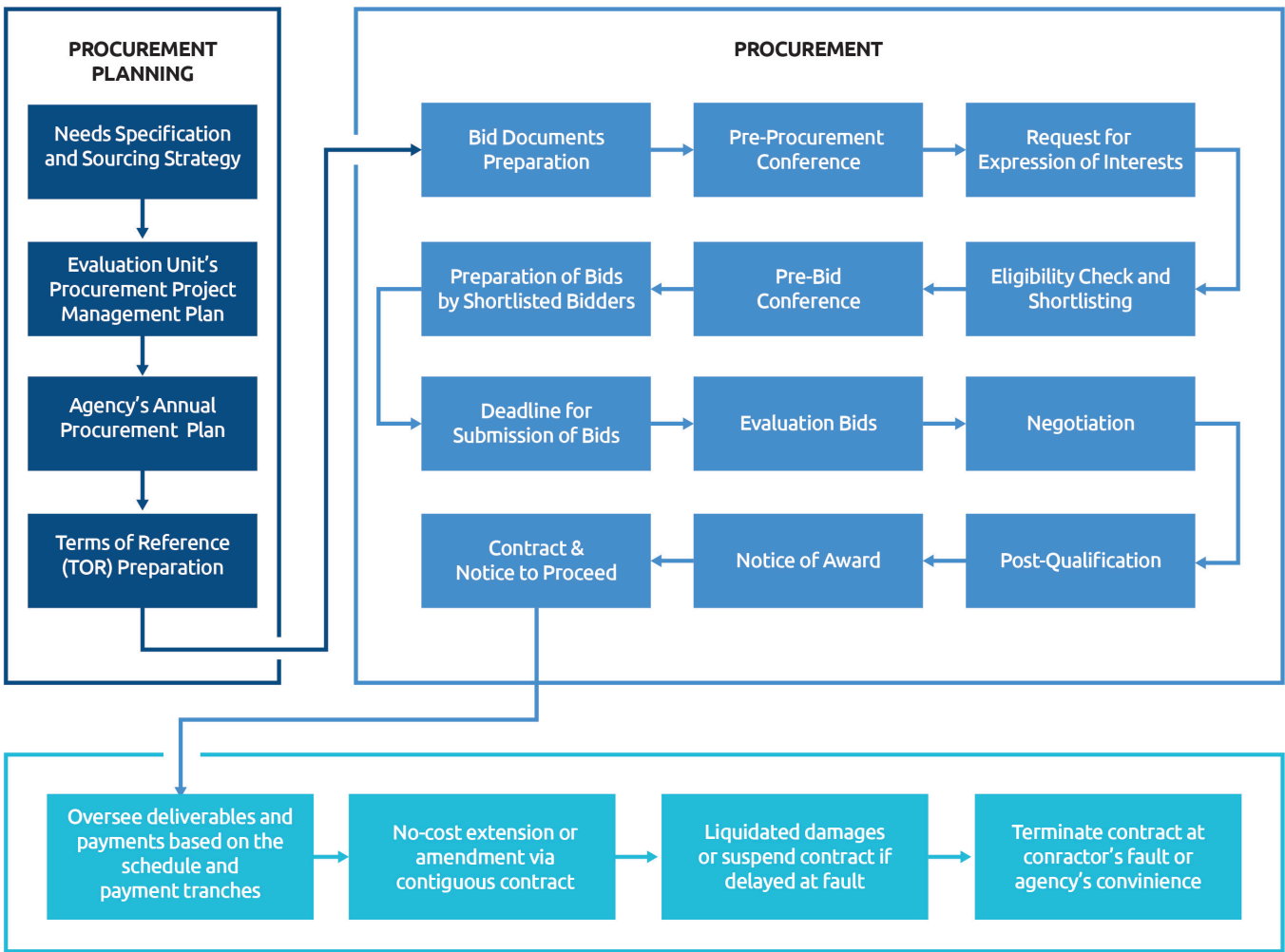
The 2016 IRR and Volume 4 of the General Procurement Manual provide detailed guidelines on procuring consulting services. The following summarize key points from the said regulations that are relevant to procuring evaluations, and to provide other guidance and suggestions, that could aid preparations for the procurement process.

## D.1 Procurement planning and strategy

(based on GPRA 2016 IRR Section 7)

A. As with all procurement undertakings, an **evaluation contract** shall be within the approved budget as indicated in the agency's Annual Procurement Plan (APP) and consistent with the agency's approved budget.

Figure D1: Procedures for Procuring Evaluation Services (Consulting)



Source: GPRA 2016 IRR

The 2016 IRR requires the preparation of two versions of the APP: the indicative APP that is submitted together with the agency's proposed budget, and the final APP after the budget is approved; with supplemental APPs or modifications every six (6) months or as required.

B. Before an evaluation project is included in the agency's APP, the **agency's evaluation unit**—which serves as the procurement management office for all evaluation projects—must collaborate with the evaluation commissioners (or the program or project management unit) in defining its scope and methodology, developing its Terms of Reference (ToR), and determining the budget estimate and source of funding for the evaluation project.

C. **The sourcing strategy**—including the appropriate mode of procurement, timing of procurement, cost structure, and risks identification—shall also be discussed and agreed upon between the evaluation unit and the relevant evaluation commissioner and in consultation with the agency’s procurement unit.

Procurement strategy shall consider the supply market, including the availability of evaluation consultants and other service providers, prevailing costs or the cost of similar past contracts, and the general macroeconomic environment.

D. **An accomplished Evaluation Plan** is necessary to determining the sourcing strategy. Short of having a fully-developed ToR, the evaluation plan identifies the services required, the outputs to be produced, and the type of capabilities needed to undertake the evaluation design: information that is necessary to costing the evaluation and identifying the best procurement approach.

E. **A key decision point** is in selecting between firms or individual(s) to undertake the evaluation, as either modality poses their respective advantages and risks.

F. **The Approved Budget for the Contract (ABC)** should be set based not solely on the projected acquisition value of the contract (i.e., the cost of the professional services) but also the total cost for the undertaking. The following must be considered:

- The total amount of the ABC shall be composed of remuneration costs, reimbursable costs, and contingency:
  - Remuneration includes salaries of the professional staff (basic salaries and social charges), overhead costs which shall not exceed 120 percent of the basic salary, and management fee that is within 15 percent of total remuneration. Payment is tied to the submission and acceptance of outputs.
  - Reimbursable costs are all other costs associated with the execution of the services, which may be based on agreed fixed rates (e.g., per diems and allowances) or on actual costs (e.g., airfare) that are supported

by invoices or other supporting documents. These may be reimbursed together with the tranche payments or as soon as activities are undertaken.

- Contingency, if provided in the contract, may be made within 5 percent of the total contract cost. For evaluation contracts, contingency is not typically required but may be for additional work that are within the general scope of the contract but are not specified as outputs (e.g., additional survey sample) or for costs that exceed original estimates (e.g., sudden increase in transport prices).
- In defining the cost structure, including terms of payment, it is useful to think in advance of their implications on contract management. Some guidance:
  - The type of contract will influence the administrative ease and degree of control during contract management. Lump sum contracts, where payments are tied to outputs, are preferred over the other types (see section D.4).
  - Consultancy services are fixed price contracts. Extensions can only be at no added cost. For additional works that may be beyond the provisioned contingency, a separate procurement process—whether competitive or through alternative modes (see section D.3)—will be required.
  - Typically, the first payment made under an evaluation contract is tied to the submission and acceptance of an inception report and may be equivalent to about 10 to 20 percent of the total contract cost. When requested by the consultant, advance payment for mobilization may be made but not more than 15 percent of total cost and only with a consultant’s submission of an irrevocable standby letter of credit.
  - It is recommended that the final tranche payment—to be paid until satisfactory submission and acceptance of the final output—be no less than 20 percent of the total contract cost. Imposing of “retention money” on progress billings, as in goods and infrastructure, is not required for consulting services.
  - Subcontracting may only be up to a maximum of 20 percent of the services.

G. **The procurement plan and strategy** will be reflected into the evaluation unit's Procurement Project Management Plan (PPMP). The PPMP feeds into the indicative and final APP of the agency. The PPMP contains the general description of the evaluation, objective, quantity, procurement method, schedule, and ABC.

H. **The Evaluation Plan**—which indicates indicative scope, methodology, schedule and costing—should accompany the indicative PPMP. At the latest, the TOR shall be submitted for BAC review and consideration with sufficient time and allowance prior to the targeted award date.

It is recommended that early procurement be undertaken for an evaluation contract so that the evaluation is undertaken at the start of the year or as soon as the new budget is passed. Refer to Annex A for an illustrative timeline; and to the template Evaluation Plan (template 2) and TOR (template 4).

I. Though currently not required by the **government procurement regulations**, it is advisable for an evaluation manager to prepare a risk register or matrix for the procurement and contract management as this helps in planning ahead for such risks and, when relevant, in incorporating risk mitigation measures in the TOR and contract itself

This risk register may be part of the overall risk matrix for the evaluation (see Sections 4.4 and 5.2 of the NEPF Guidelines).

Table D.1: Selecting Between Individuals and Firms

	Individual	Firm
Advantages	<ul style="list-style-type: none"><li>■ Enables agency to select the most appropriate individual(s) based on specialization and experience</li><li>■ Better control of operating costs (e.g., meeting arrangements) and total costs may be less expensive</li><li>■ Faster procurement and less requirements on bidders</li><li>■ Opportunity for capacity development through mentoring and knowledge exchange</li><li>■ May be more flexible to changes to the TOR and other arrangements</li></ul>	<ul style="list-style-type: none"><li>■ Firms can tap into their internal staff and consultants' rosters for the required specializations.</li><li>■ Allows agency to leverage the firms' established expertise and internal quality assurance mechanisms</li><li>■ Enables agency to reduce risk burden, e.g., burden on replacing underperforming members on firm</li><li>■ Enables agency to pass on many administrative and logistical burdens to the firm (e.g., arranging for travel and meetings).</li></ul>
Disadvantages	<ul style="list-style-type: none"><li>■ Administrative burden and costs (e.g., arranging meetings, travels, etc.) falls on the agency</li><li>■ Relationship among a team of individual consultants, including conflicts, will need to be managed</li><li>■ Risk burden largely falls on the agency, e.g., security of team during field work in conflict-prone areas</li><li>■ Burden of replacing members falls on agency, e.g., if a team member leaves or underperforms</li></ul>	<ul style="list-style-type: none"><li>■ Less control of administrative costs (if a "lump sum" fee is proposed by firm) and may be more expensive.</li><li>■ More lengthy and cumbersome bidding process, and more requirements on the bidders</li><li>■ Team members may tend to have similar approaches and perspectives</li><li>■ Risk of subcontracting by firms</li><li>■ Risk of hiring firms which are "shells" (i.e., no in-house capacity and resorts to hiring freelancers)</li></ul>
Other Considerations	<ul style="list-style-type: none"><li>■ Schedule of payments if based on deliverables or daily/monthly fee</li><li>■ Connection between consultants' outputs and their sequencing (e.g., for purposes of payment)</li></ul>	<ul style="list-style-type: none"><li>■ Appropriate cost structure including proportion of management fees</li><li>■ Insurance for team members especially for risky field work</li><li>■ Joint ventures and partnerships</li></ul>

Adapted from UNDP (2018) with modifications

# D.2 Competitive procurement for evaluation consultants (firms and individuals)

(based on GPRA 2016 IRR Sections 20 to 37, Annex B and Annex C)

- A. The competitive procurement process for evaluations is typically undertaken as a procurement for consulting services, which entails two stages: the shortlisting of eligible bidders based on experience and workload; and the evaluation of technical and financial proposals of shortlisted bidders.
- B. Excluding planning and the preparation of bid documents, procurement takes a minimum of 36 days and a maximum of 180 days from the moment the Request for Expression of Interest (REOI) is advertised to the issuance of the Notice to Proceed (see Figure D.1 and Table D.2).

Table D.2: Timeline of procurement process for an evaluation

Activity	Calendar Days Required		Timeline (Days)	
	Minimum	Maximum	Minimum	Maximum
Pre-Procurement Conference <sup>1</sup>	1	1	0	0
Advertisement of REOI <sup>2</sup>	7	7	1 to 7	1 to 7
Eligibility Check & Shortlisting	1	20	8	8 to 27
Preparation of Bids	17	53	9 to 26	28 to 81
Pre-Bid Conference <sup>1,3</sup>	1	1	15	34 to 70
Deadline of Submission of Bids	1	1	27	82
Bid Evaluation <sup>4</sup>	1	21	28	83 to 103
Approval of Ranking	1	2	29	104 to 105
Notification for Negotiation	1	3	30	106 to 108
Negotiation	1	10	31	109 to 118
Post-Qualification	2	30	32 to 33	119 to 148
Issuance of Notice of Award	1	15	34	149 to 163

Contract Preparation & Signing <sup>5</sup>	1	10	35	164 to 173
Issuance of Notice to Proceed	1	7	36	174 to 180

Adapted from GPRA 2016 IRR Annex C

Notes: 1) optional for ABC of PHP 1 million and below; 2) newspaper advertisement required for ABC of PHP 15 million and above; 3) not earlier than 7 calendar-days from approval of shortlisting and with at least 12 calendar days before deadline of bid submissions; 4) bidders may be required to make an oral presentation within 15 calendar days after the deadline of submissions; 5) additional 30 to 40 days may be taken if approval of higher authority is needed.

- C. The above-mentioned procurement timeline is for both individuals and firms, and the agency must optimize the process to secure a consultant at the right time, with the equally-important objective of being able to encourage the best proposal possible. The following are pointers to consider in planning the procurement process:

- An agency may want to achieve the minimum timeline to procure an evaluator; however, it is advisable to allow sufficient time for the bidders to prepare quality proposals, and for the bid evaluation team to review these.
- The length of time and requirements for the procurement of a firm, compared to an individual consultant, may be daunting. It must be noted, however, that procuring a team of individuals separately—and managing their contracts once awarded—may just be as cumbersome (See Table D1).
- One may be enticed to skip the pre-procurement and pre-bid conferences for contracts below PHP 1 million to reduce time. Consider, however, that such conferences allow the procuring entity to improve on the TOR and bid documents and, therefore, save on time. Unclear specifications and requirements often cause delays and problems: from the need to issue frequent bid bulletins; to receiving bids that are difficult to rate due to lack of the necessary information.

- D. Shortlisting is a crucial step in the procurement of consulting services as it determines the most eligible consultants to be invited to prepare and submit technical and financial proposals.



- Shortlisting enables the government to assess and receive the best possible proposals. First, preparing quality bids for consulting contracts requires time and resources from bidders, who may be discouraged from doing so if the chance of being selected is narrow. Filtering for the best-qualified consultants is deemed to encourage them to prepare quality bids. Second, it is time-consuming for the agency to review the technical and financial proposals for consultancy contracts and limiting the review to those of the shortlisted bidders is efficient.
- A minimum of three (3) and a maximum of seven (7) eligible consultants should be short listed. If there are less—even just one (1)—bidders which applied for and passed eligibility and shortlisting, then they shall still be considered. Having a single qualified and shortlisted bidder is not grounds for declaring failure of bidding.
- The 2016 IRR requires the following considerations for the shortlisting, with the following recommended weights according to the Generic Procurement Manual:
  - Experience of the firm or joint venture members (50 percent). This refers to the applicable experience, in terms of years in business and previous engagements or contracts, with respect to the project at hand. For individual consultants, this criterion is taken out in favor of the below, which should then include past similar projects in his/her statement of relevant experience.
  - Qualification of key staff assigned to the job (30 percent). This typically is measured by the relevant experience, training, and education of the key personnel to be assigned by the firm. At the very least, this shall be the project lead, who should ideally be an incumbent or “organic” personnel (i.e., as opposed to a hired consultant or freelancer) of the firm.
  - Current workload relative to capacity (20 percent), which is usually taken as financial capacity, but should also incorporate an assessment for the ability to undertake additional works given existing contracts. The Net Financial Contracting Capacity (NFCC) is not a requirement but may be used; although simpler financial metrics, such as net working capital (current assets minus current liabilities), may be more advisable.

- The shortlisting criteria and rating scheme should be clearly stated in the ToR and bid documents and should easily be derived from the submission documents required from the bidders.

E. The evaluation scheme for assessing the proposals of shortlisted bidders may either be Quality-Based or Quality Cost-Based, which have their advantages and disadvantages:

- **Quality-Based Evaluation** considers solely the rating of technical proposals against set evaluation criteria. In this mode, the technical proposals are opened first for scrutiny and the highest rated bid is approved by the head of agency.

Quality-Based Evaluation may encourage firms to maximize their technical proposals and bid very closely to the ABC. Although not commonly used, this mode may be advisable when the requirements are complex and difficult to precisely define in the ToR, or if the agency seeks to encourage innovative proposals.

- **Quality Cost-Based Evaluation** assesses both technical and financial proposals, typically with a distribution of 80-20 percent; the latter can take a weight of a minimum of 15 to a maximum of 40 percent of the total score. To identify the highest rated bid, the technical proposals are reviewed first followed by the financial proposals of those meeting a minimum technical score (70 percent).

This mode, which is more commonly used, may be more appropriate when the agency has rigid specifications for the methodology and requirements, and seeks to achieve the most efficient market cost for delivering the said outputs.

- **Whether Quality-Based or Quality Cost-Based**, technical evaluation must be based on evidence from the documents submitted and, when required, oral presentations from the bidders. In general, the criteria are: (a) quality of personnel; (b) experience and capability of the consultant; and (c) plan of approach and methodology.

- **Quality of personnel** – how appropriate are the key staff proposed by the bidder to be assigned to the project? As in the shortlisting criteria, metrics

include length of relevant experience, specialized training, and educational attainment of all the key personnel and consultants to be assigned by the firm.

F. Experience and capability of consultant – in the case of firms, experience and capability is typically rated through the profile of the firm itself in terms of experience (years in business), leadership, and organization; and performance in previous projects of similar scope and nature, including satisfaction of clients.

- Plan of approach and methodology – the bids should be assessed based on their clarity and completeness, feasibility, innovativeness, and comprehensiveness These should be assessed by how the bidder understands the requirement of the TOR and its context, how responsive the proposals are in meeting or exceeding such requirements, and how the bidder intends to assure the quality of the outputs and mitigate the risks. Since the ToR broadly identifies the methodology and milestones, the bidders must provide the specific methodology and work plan.
- The weights assigned will depend on what the agency prioritizes for the evaluation consultant, for instance, whether the capability of the personnel is more important than the firm’s experience; or if more weight is given to the soundness of the proposed methodology compared to the capacity of the firm.
- Below is an illustration of the technical evaluation criteria and scoring for an evaluation firm. This may be modified as needed for the specific ToR:

Table D.2: Timeline of procurement process for an evaluation

	Sample 1: Balanced Rating	Sample 2: Emphasis on Methodology	Sample 3: Emphasis on Capacity
I. Quality of Evaluation Team Overall: composition of the team & allocation of tasks:			

<div>Lead Evaluator<ul style="list-style-type: none"><li>■ Experience</li><li>■ Training</li><li>■ Education</li></ul></div> <div>Evaluator 1<ul style="list-style-type: none"><li>■ Experience</li><li>■ Training</li><li>■ Education</li></ul></div> <div>Evaluator 2<ul style="list-style-type: none"><li>■ Experience</li><li>■ Training</li><li>■ Education</li></ul></div>	40	30	50
<div>II. Experience and Capability of Consultant (Firm):<ul style="list-style-type: none"><li>■ Capacity of the firm in terms of experience, organization and leadership</li><li>■ Performance in previous projects, quality of outputs, and relationship with past clients</li><li>■ Financial capacity</li></ul></div>	30	10	30
<div>III. Plan of Approach and Methodology<ul style="list-style-type: none"><li>■ Clarity and completeness of proposed approach and methodology to</li></ul></div>	30	60	20

<div>respond to the TOR requirements</div> <div><div><div>■ Innovativeness of proposed methodologies to respond to the evaluation study objectives/questions</div><div>■ Feasibility of the work plan given the approach and methodology and the risks to the project.</div></div></div>			
--	--	--	--

G. Negotiations are then undertaken with the consultant with the highest (or single) rated bid. Negotiations cover the discussion and clarification of the ToR and scope of services, finalization of methodology and work program, consideration of personnel compensation and manning schedule, data and other services required from the agency, the financial proposal, and contract provisions.

For evaluations, it is important to level off on the specificities of the methodology to be employed and the work program, including the data, approvals (e.g., for surveys), and other requirements to be secured by the agency commissioning the evaluation for the consultant.

H. After successful negotiations, post-qualification is undertaken to determine whether the highest (or single) rated bidder is responsive to all the requirements and conditions for eligibility. In post-qualification, a bidder’s compliance with the legal, technical, and financial requirements of the tender is verified.

At this stage, it is crucial to ascertain that the consultant is in good standing and has not been blacklisted by any government entity. The highest (or single) rated and responsive bid is then declared after a successful post-qualification. The issuance of the notice of award, contract signing, and issuance of notice to proceed follow.

I. A failed bidding occurs when no bids are received, the bids are all declared ineligible or fail post-qualification, there are no successful negotiations, or the highest rated and responsive bidder does not accept the award.

Before undertaking a re-bid or a negotiated bid (after two failed biddings, see D.3. below), it is useful to understand why such a failure took place.

- Were the qualifications and technical requirements too steep to be provided by the market?
- Are there unnecessary documentary submissions and other requirements?
- Were the criteria too restrictive or too vague? Is the ABC enough for the requirement?

The TOR and bid documents may be adjusted before a re-bid or a negotiation after two failed biddings, taking note that for the latter a maximum ABC increase of 20 percent is allowed.

# D.3 Alternative procurement methods applicable to evaluations

(based on GPRA 2016 IRR Sections 48 to 54 and Annex H)

## Competitive Procurement

**Competitive procurement** is the default mode of procurement under the 2016 GPRA. It is also appropriate for evaluation projects as these tend to be of high value (i.e., more than PHP 1 million). Competitive procurement ensures that only the most qualified suppliers, with the most responsive and innovative proposals, are selected.

## Alternative Procurement

However, there may be instances where **alternative procurement** methods are more appropriate: whether to optimize efficiencies (considering that competitive procurement takes time), leverage partnerships and government capabilities, or when warranted by circumstances. In all cases, the use of alternative methods must be approved by the head of the procuring entity and indicated in the agency’s APP.

## Negotiated Procurement

Among the alternative procurement methods, **negotiated procurement** is the most relevant for consulting services for evaluations. Under negotiated procurement, a procuring entity directly negotiates a contract with a technically, legally, and financially capable consultant. Specifically, the negotiated method may be applied for the following cases for evaluation projects: a) two failed biddings; b) take-over of contracts; c) adjacent or contiguous; d) agency-to-agency; e) scientific or scholarly work; f) highly technical consultants; g) small value procurement; and h) procurement from United Nations (UN) agencies and other international organizations.

Additionally, limited source bidding and exclusive procurement for non-government organizations (NGOs) may in the future be possible, but are not feasible at present.



**A. Two Failed Biddings** – when a bid fails for the second time, the agency’s BAC invites at least three (3) consultants for negotiations and to submit a best offer based on the final technical and financial requirements agreed during the negotiations. The single or highest rated and responsive proposal is then recommended for

award. It must be noted that negotiated procurement after two failed biddings should still be posted on the Philippine Government Electronic Procurement System (PhilGEPS).



**B. Take-Over of Contracts** – for evaluation projects, this may be considered when a contract previously-awarded through competitive bidding has been rescinded or terminated for causes provided for in the contract and existing laws. Here, the BAC post-qualifies and negotiates with the next highest-rated bidder at the said bidder’s own original bid price applicable to the remaining work.



**C. Adjacent or Contiguous**—for consulting services including evaluations, an adjacent or contiguous contract pertains to one which requires subject matters, outputs, and/or deliverables that are closely linked or related to those in an ongoing contract.

Adjacent or contiguous contracts may be resorted to only if a) the original contract was competitively procured, b) the original and the adjacent or contiguous contracts have related scopes of work; c) the adjacent or contiguous contract is within the contracting capacity of the consultant; d) the consultant uses the same or lower unit prices; e) the amount of the contiguous or adjacent work does not exceed the amount of the ongoing project; and f) the consultant has no negative slippage or delay.

The adjacent or contiguous contract should be negotiated with the consultant of the existing contract before the latter expires. Here are examples of when an adjacent or contiguous contract may be used in the context of evaluations:

- Evaluation of a related program, project, or component that has similar mechanics as the one originally evaluated, and this would entail minimal changes to the evaluation methodology and data collection;
- Evaluation of the same program or project but in different locations, but essentially the same evaluation methodology and data collection will be used;
- The coverage of data collection—e.g., survey (sample size), key informant interviews (number) and focus group discussions (number and size)—need to be increased, provided that unit costs remain the same or are lower;



**D. Agency-to-Agency** – for evaluations, an agency may enter into a memorandum of agreement (MOA) with a public research institution such as the Philippine Institute for Development Studies (PIDS), a State University or College (SUC), or other government body with the mandate and capacity for research and evaluation.

- Agency-to-agency procurement for evaluations may be undertaken if: it is justified to be more efficient and economical to the government, the servicing agency has the mandate to deliver the said service, the servicing agency has the absorptive capacity for the project, and the agency owns or has access to the tools to undertake the project.
- Sub-contracting by the servicing agency is not allowed. Moreover, all procurements to be undertaken by the servicing agency shall continue to be governed by the GPRA and its IRR.



**E. Scientific, scholarly, or artistic work** – this method of procurement allows for contracting with a particular supplier, contractor, or consultant for scientific, academic, scholarly work, or research, as well as subscription to scientific, technical, economic, business or trade journal, or statistical publications and references.

- This may be relevant, for example, in tapping an academician scholar who has already established research outputs, research methodologies, and data relative to the program or project to be evaluated, and the agency seeks to leverage the knowledge and techniques already developed by the said academician.
- Prior to undertaking procurement under this method, the end-user unit of the agency shall conduct a market study which determines probable sources of supply and confirms that the consultant could undertake the project at more advantageous terms. For evaluations, this may entail a comprehensive literature review to determine authors who have done work in a relevant research area.



**F. Highly technical consultants** – this pertains to the hiring of individual consultants to undertake work that is either highly technical or proprietary, or; primarily confidential and policy determining, where trust and confidence are the primary considerations for selecting the consultant.

- For evaluations, this mode may be appropriate for the hiring of a subject-matter expert to guide the design and quality assurance of an evaluation, to advise the agency's capacity development for monitoring and evaluation, to conduct a confidential rapid assessment on a program, among other applicable work.

It is not recommended for the hiring a team of individual consultants that will form an evaluation team. While the lead evaluator may be proven to be a sought-after “highly technical” policy adviser, for example, the rest of the evaluation team (e.g., analysts or associates) may readily be available in the supply market.

- A highly technical consultancy may only be for at most six (6) months, renewable at the option of the head of procuring entity but within his or her term of office.



**G. Small value procurement** – this pertains to the procurement by a national agency of consulting services with ABC below PHP 1 million. Under this method, the BAC prepares a request for proposal to be sent to at least three (3) consultants of known qualification and, for projects with ABC above PHP 50,000, posted on PhilGEPS.

- This modality may be appropriate for the hiring of a consultant, likely one (1) individual but may also be a firm, for a smaller-scale evaluation project or related work, e.g., a rapid assessment a midterm or termination of a project, an evaluability assessment, a theory of change workshop, among others.
- While small-value procurement may be used for the hiring of a team of individual evaluators, the following must be considered. First, the procuring entity must justify that this does not result in the splitting of contracts (see #14 below). Second, the procuring entity absorbs the administrative and logistical burdens of carrying-out the evaluation. Third, while the agency may be able to hire the “best of the best” among evaluators, there are also risks to hiring individuals who have not worked together.

Ultimately, this strategy may be pursued if it is justified to be more efficient and economical to the government than getting a firm.





**H. UN Agencies and Other International Organizations** – for evaluations, procuring from such specialized agencies of the UN and other international organizations and financing institutions may be applicable if the requirement is for consulting services that involve advanced technologies, techniques, and innovations which are not locally available as certified by the head of procuring entity.

1. **On NGO Participation** – NGOs including civil society organizations (CSOs), academic and research institutions, and other not-for-profit entities may participate in the procurement of evaluations and compete against commercial consulting firms. However, the use of negotiated procurement method for NGO Participation—which is a process exclusively for NGOs and other non-profits—may not yet be practicable at present as the modality is only relevant when an appropriation law earmarks funds or projects specifically for NGO contracting.
2. **On Limited Source Bidding** – also known as selective bidding, this alternative method of procurement entails the pre-selection of consultants with known experience and proven capability. This is resorted to for highly specialized consulting services, where only a few consultants are known to be available and resorting to the normal competitive process will likely not result in additional participants.

This method is not yet feasible as it requires a relevant government authority to maintain a list of such consultants and updates the same periodically. Limited source bidding generally follows the competitive bidding process, except that the request for expressions of interest is not advertised publicly but only sent to the consultants in the said roster.

3. **No Splitting.** In resorting to the above-mentioned negotiated procurement methods, the procuring entity must ensure that such does not result in the splitting of government contracts. Splitting means the breaking up of contracts into smaller amounts or into artificial implementation phases to circumvent the competitive bidding and the transparency and accountability requirements of the GPRA.
- Splitting happens if the scope of work is clear from the beginning and yet the agency chooses to subdivide the work into phases artificially for the sake of

skirting thresholds. For instance, a PHP 15-million evaluation contract is split artificially into two phases to skirt the advertising requirement; or a PHP 5-million project is divided into six phases so that small-value procurement can be used.

These practices are unethical for circumventing competitive bidding, and especially worse if the split contracts end up in the hands of a single consultant. Splitting also occurs if an agency hires individual evaluators that belong to the same firm as individual consultants, rather than undertaking a competitive tender for a firm.

- Splitting, however, is different from sound procurement strategies that maximize efficiency and economy. For instance, the hiring of a team of individual consultants may be considered if it is found from the market research that the needed specializations cannot be easily found in a single consulting firm.

The disadvantage here is that the agency would absorb the administrative burdens that could have been passed on to a firm. It may also be sound to procure multi-phase evaluations as separate contracts, where each wave is years apart and where the scope of work and/or funding for the future waves are not yet clear.

Hiring an individual consultant for preparatory work, such as to prepare a detailed evaluation design, prior to bidding out the full evaluation to a firm is advisable.

## D.4 Contract Management

(based on GPRA 2016 IRR Sections 39-42, 59-62, and 68 and Annex F and Annex I)

- A. Whether the procurement is done competitively or through alternative methods, it is advantageous to think ahead and plan for the implementation of the proposed evaluation and integrate these considerations in the evaluation contract itself.

These considerations include the cost structure of the contract; the evaluation manager's levers in managing an evaluator's contract; and available "nuclear" options in case the contract does not go well as intended.



B. In defining the scope of work and outputs of the evaluation project, it is useful to assess the implications of the payment structure on contract management. The contract types that are relevant to evaluations are lump sum and time-based contracts.

- Typically, evaluation contracts are lump sum contracts. Lump sum contracts are advisable for their ease of administration as payments are tied to the submission and acceptance of outputs. For such contracts, the scope of work and timeline of deliverables shall be specified in the TOR.

Reimbursable costs may or may not be specified. With the former, the agency has better control of costs but assumes a lot of administrative burden in terms of reviewing and processing reimbursements. It would be easier if all costs are incorporated in the remuneration.

- Time-based contracts are much more tedious to administer but provides flexibility when the scope and length of the services is difficult to define at the onset. Payments for such contracts are based on agreed rates for the staff and reimbursables.

The contract should specify the maximum total payments, including contingency. For evaluations, these may be relevant for a “design-and-build”-like arrangement, where the consultant is tasked to first undertake an evaluability assessment or process evaluation before defining and undertaking a more rigorous evaluation of outcomes or impact. It may also be applicable for an “on call” arrangement with a consultant who is tasked to conduct quality assurance or peer review of ongoing evaluations.

- To recall, all consultancy services are fixed price contracts and any extension of contract time shall not involve additional costs. This applies whether the contract structure is a lump-sum or time-based contract.

C. An evaluation manager must establish a protocol for monitoring and managing the progress of the evaluation consultant and for assuring the quality of outputs. To avoid conflicts and misunderstanding during contract implementation, protocols and parameters must be clearly built into the TOR and contract terms. These include the review and approvals required prior to accepting and paying outputs, who these

reviewers are, and the quality standards by which the outputs will be reviewed, as defined in the NEPF Guidelines. The GPRA 2016 IRR require the following:

- The winning bidder must post a performance security prior to signing of the contract. It may be in the form of cash or a cashier’s/ manager’s check, a bank draft/guarantee or irrevocable letter of credit that is worth five percent of the total contract price; or a surety bond of 30 percent of the contract price.

This is required for competitive bidding, two failed biddings, take-over contract, adjacent or contiguous contracts; and may be considered for small value procurement.

- Warranty security is not required for consulting services except when these pertain to detailed engineering design and construction supervision. Similarly, “retention money,” which may be imposed for goods or infrastructure, is not required for consulting services. In other words, firms hired to perform evaluations are only required to file a performance security but not warranty security or retention money.

The additional safeguard that could be built into an evaluation contract is that the final payment should not be less than twenty percent of the total contract price, paid only once the final evaluation report and other products are submitted and accepted.

- No replacement of personnel can be allowed until after 50 percent of the personnel’s man-months have been served, except for justifiable reasons that may be determined by the BAC, such as death, illness or incapacity, or resignation. Any unjustified replacement shall be fined an amount that is at least 50 percent of the total basic rate of the said personnel for the duration of the engagement.
- Subcontracting may be allowed to an extent subject to the review and approval of the procuring entity, including the subcontracting agreements and amendments thereto.

The subcontracted work should not be a material portion of the project and should not exceed 20 percent of the total project cost.

The consultant retains full responsibility for the project and the quality of outputs.

D. In case an evaluation consultant fails to perform, the evaluation manager has the following levers provided by the GPRA and 2016 IRR, which are reiterated in the bidding documents and TOR:

- An evaluation manager may grant time extensions based on justifiable reasons, for example, if data gathering is hampered by the unavailability of information or required persons.

However, for delays and incompletion due to the fault of the consultant, liquidated damages may be imposed that is equivalent to one-tenth (1/10) of one (1) percent of the cost of the unperformed portion for every day of delay. Once the liquidated damages reach 10 percent of the contract amount, the contract may rescind or terminate the contract, without prejudice to other courses of action or remedies such as forfeiture of the performance security.

- Suspension of work and of payments is another lever. The procuring entity may suspend payments to it if the consultant fails to perform any of its obligations due to its own fault or due to force majeure, until such time that the failure is remedied but within 30 days from notice.

The consultant can also suspend work affected by the procurement entity's failure to supply the data, facilities, and other agreed-upon requirements; peace and order conditions; and delays in the payment of consultant's claim for billing beyond 45 days from filing.

- The procuring entity may terminate a contract, in part or in full, under the following circumstances. Annex "I" of the 2016 IRR provides more guidance:



**Default:** when, outside of force majeure, the consultant fails to perform the outputs within the periods in the contract or within any extension granted, such as when the liquidated damages reach 10 percent of the total contract amount; or if it fails to perform any other obligation under the contract.



**Convenience:** the procuring entity may terminate the whole or parts of the contract at any time for its convenience, if conditions deem the project economically, financially, or technically impractical or unnecessary, such as fortuitous events and changes in government policy.



**Insolvency:** if the consultant is declared bankrupt or insolvent as determined with finality by a court of competent jurisdiction.



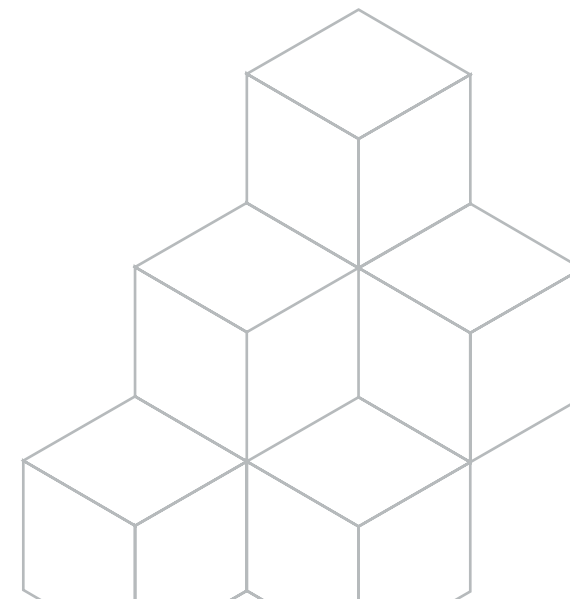
**Unlawful Acts** – if it is determined prima facie that the consultant has engaged in unlawful deeds and behaviors relative to contract acquisition and implementation. These include corrupt, fraudulent, collusive, or coercive practices; using forged documents or adulterated methods; among others.

The consultant may also terminate the contract with the procuring entity if the latter is in material breach of its obligations pursuant to the contract and has not remedied the same within 60 days of notice of the breach.



# Templates

- Evaluation Agenda
- Evaluation Plan
- Evaluability Checklist
- Costing an Evaluation
- Terms of Reference (TOR)
- Inception Report Outline
  - Includes Evaluation Matrix
- Evaluation Report Outline
- Evaluation Report Checklist
- Management Response
  - Includes Implementation Plan
- Quality Assurance Checklist



# Evaluation Agenda

PDP Out-come	Program/Project to be Evaluated	Program/Project Implementation Unit	Type of Evaluation (Function & Method)	Estimated cost (PHP)	Target Date of Completion	Source of Funding	Planned Start Date

Note: this Evaluation Agenda should be supported by initial draft Evaluation Plans as attachments

## Evaluation Plan

I. Program/Project Information	
1. Name of Program/Project	(if thematic evaluation, identify the theme and the component programs, if any, that will be evaluated)
2. Program/Project Location(s)	
3. Program/Project Duration	
4. Program/Project Status	<div><input type="checkbox"/> New</div> <div><input type="checkbox"/> On-going</div> <div><input type="checkbox"/> Completed</div> <div><b>Budgeting classification:</b></div> <div><input type="checkbox"/> Tier 1</div> <div><input type="checkbox"/> Tier 2</div> <div><b>Details:</b> (Provide additional descriptions e.g. pilot experimental stage, ongoing but delayed or advanced, winding down/for closure, etc.)</div>
5. Lead Government Agency	
6. Other Government Agencies	
7. Other Implementing Partners	
8. Program/Project Intent and Rationale (discuss the problem that the program is trying to solve in 2-3 paragraphs, including policy pronouncements)	
9. Alignment with the Philippine Development Goals (identify PDP Chapters relating to the program)	
10. Alignment with the Sustainable Development Goals (identify SDGs relating to the program)	

<p><b>11. Results Framework</b> (attach logical framework, theory of change, or other framework to the evaluation plan, ideally indicating the output and outcome statements, indicators, and targets aligned with the PDP Results Matrixes and/or SDG indicators)</p> <p><input type="checkbox"/> Is the results map logical framework and/or Theory of Change clearly defined?</p> <p><input type="checkbox"/> Are the indicators clearly stated in the results framework?</p>
<p><b>12. Program/Project Financing</b> (describe the program financing strategy, sources, and aggregate amounts for the whole duration of the program, with description of trends over time. Attach filled-out template to the evaluation plan)</p>
<p><b>II. Information on the Evaluation</b></p>
<p><b>13. Key Evaluation Stakeholders</b> (apart from those already identified in #4, 5, and 6, list down the other stakeholders involved in or affected by the program and/or who are crucial to the success of the evaluation.)</p> <p><input type="checkbox"/> Is the planned evaluation relevant and useful to key stakeholders?</p> <p><input type="checkbox"/> Are the stakeholders committed to support the evaluation?</p>
<p><b>14. Purpose of the Evaluation</b> (2 paragraphs or set of bullet points that identify i) the research objectives for the study, and ii) the policy goals of the evaluation, e.g., to influence programming and resource allocation.)</p>
<p><b>15. Key Evaluation Questions</b> (questions that measure effectiveness, appropriateness and efficiency.)</p> <p><input type="checkbox"/> Are the evaluation questions feasible given the: (1) project design, (2) data availability, and (3) resources available?</p> <p><input type="checkbox"/> Are the evaluation questions of interest to key stakeholders?</p>
<p><b>16. Date</b> (2 paragraphs/set of bullets that i) list down the data needed, including baselines, and initial comments on availability, and ii) describe, the data collection methods to be used to measure results.)</p> <p><input type="checkbox"/> Is there sufficient data collected to answer the evaluation questions?</p>

<p><input type="checkbox"/> Is there sufficient data disaggregation (e.g. age, sex, disability, ethnicity, migratory status, and geographic location where relevant)?</p> <p><input type="checkbox"/> If not, are there plans and/or means available to collect and disaggregate the data?</p>
<p><b>17. Relevant Literature</b> (Initial list of related literature, beginning with existing evaluation studies and followed by other studies and relevant references. If more than 10, attach as a separate document.)</p>
<p><b>18. Risk Identification and Analysis</b> (1-2 paragraphs or set of bullets identifying the key risks facing the evaluation, and initial measures to mitigate or manage these. Attach detailed risk log if already available.)</p> <p><input type="checkbox"/> Will physical, political, social, and economic factors allow for an effective conduct and use of evaluation as envisaged?</p>
<p><b>19. Evaluation Timeline</b> (Indicate overall timeframe and up to five key milestones. Attach detailed timetable or Gannt chart for the evaluation project, if already available.)</p> <p><input type="checkbox"/> Is there sufficient time for the evaluation?</p>
<p><b>20. Evaluator Profile</b> (Indicate type of provider and rationale for the choice. E.g., in-house or contracted out? Firm or individuals? Academic institution or think tank? What are the necessary competencies, skills and subject matter specialization are needed? Do not indicate preferred consultant or entity unless the procurement modality allows for direct contracting.)</p> <p><input type="checkbox"/> Are there available service providers?</p>
<p><b>21. Indicative Resource Requirements</b> (in PhP, indicating both costs for the evaluation consultant or firm and for evaluation management, e.g., activities to be sponsored by project management team)</p> <p><input type="checkbox"/> Are there sufficient human resources?</p> <p><input type="checkbox"/> Are there sufficient financial resources?</p> <p><input type="checkbox"/> Are there available service providers?</p>

Attachment: Program Funding				
Period of Funding: (e.g. 1 July 2014–30 June 2017)				
Total Program Funding: (PhP Million)		Period of Funding: (e.g. 1 July 2014 – 30 June 2017)		
Annual Program Funding Breakdown: (Indicate N.A. if not applicable, or TBD if not yet known at the moment)				
Year	Government Funding Sources (e.g., GAA, GOCC)	Total Government Funding (PhP Million, Actual)	Other Funding Sources (e.g., ODA, PPP, private sector)	Total Other Funding (PhP Million, Actual)
Implementation Modality: (describe the contracting modality, indicating if largely implemented by administration or by contract, and if implementation is transferred by the main implementing agency to another agency, e.g., DepEd transferred to DPWH, who in turn procure contractors under RA 9184)				
Top Program Contractors: (list the top 5 and the total amount of contracts awarded to these from the program. Include only the contractors and not implementing partners which shall be indicated above)				
	Contractors	Total Awarded Contracts (PHP Million)		

	Contractors	Total Awarded Contracts (PHP Million)

Does the Program Have Allocations for Evaluations? Yes ☐ No ☐

(If yes, indicate aggregate amount and describe the activities. Attach costed monitoring and evaluation plan or other breakdown of the expenditures for monitoring and evaluation activities)



# Outcome Evaluation Workshop

## Evaluability Assessment Tool

### Checklist for Evaluability Assessment

EAs include information on three topics:

- 1. Project design
- 2. Availability (and quality) of information
- 3. Institutional context.

Please circle the appropriate score: 3 (MEETS REQUIREMENT); 2 (APPROACHES REQUIREMENT); 1 (NEEDS IMPROVEMENT)

## 1. Project Design

Components	Evaluability Assessment Questions	A	B	C	D	Outcome (A to D Average)	Notes/Remarks
1.1 As described in Theory of Change/ Logical Framework/ Narrative in prodoc/ UNDAF/CPAP	CLARITY: Are the <b>long-term impact and outcomes</b> clearly identified? Are the proposed steps towards achieving these outcomes clearly defined?	3	3	3	3		
	RELEVANCE: Is the <b>intended beneficiary group clearly identified</b> ? Is the project objective clearly relevant to the needs of the target group, as identified by any form of situation analysis, baseline study, or other evidence and argument?	3	3	3	3		
	CONTEXTUALIZED: Have <b>assumptions</b> about the roles of other actors outside the project been made explicit (both enablers and constraints)?	3	3	3	3		
	PLAUSIBLE: Is there a continuous causal chain, connecting the intervening projects with the final impact of concern? Is it likely that the <b>project goals</b> could be achieved, given the planned interventions, within the project lifespan?	3	3	3	3		
	AGREEMENT: Are there areas where <b>stakeholders</b> diverge on project objectives or emphasize specific aspects differently? Will these differences affect the way that project is evaluated?	3	3	3	3		
1.2 Quality of Theory of Change/ Logical Framework/ M&E Framework	CLARITY: Does a <b>theory of change/ logical framework</b> exist in which implementing organisations are clearly located and there is no discontinuity between inputs and impacts? —If not one has to be developed with the project team/ stakeholders involved as a first step in the evaluation.	3	3	3	3		
	LOGICAL: Are <b>inputs, outputs, outcomes, impacts</b> (short- and long-terms) clearly identified, and are the proposed relationships between these towards achieving impacts clearly defined?	3	3	3	3		
	VALIDITY: Are there <b>valid indicators</b> (broadly defined) for each expected event (output, outcome and impact levels)? I.e. will they capture what is expected to happen?	3	3	3	3		
	RELIABILITY: Are they <b>reliable indicators</b> ? I.e. will observations by different observers find the same thing?	3	3	3	3		
	TESTABLE: Is it possible to identify which linkages in the <b>Theory of Change</b> will be most critical to the success of the project, and thus should be the focus of evaluation questions?	3	3	3	3		

Components	Evaluability Assessment Questions				A	B	C	D	Outcome (A to D Average)	Notes/Remarks
1.3 Interference?		CONSISTENT: Is there consistency in the way the Theory of Change is described across various project documents (Design, M&E plans, work plans, progress reports, etc.)	3	3	3	3	3	3		
		STRATEGIC: Are the goals and activities aligned with current policy objectives viz. a vis. Gender and inclusion? Is the project strategy gender informed?	3	3	3	3	3	3		
		ATTRIBUTION: <b>Are other similar projects taking place via other organisations</b> in the same geographical area that could make it difficult to disentangle the impact of the project in question from other simultaneously ongoing projects?	3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		
			2	2	2	2	2	2		
			1	1	1	1	1	1		
			3	3	3	3	3	3		

Components	Evaluability Assessment Questions					Outcome (A to D Average)			Notes/Remarks
	A	B	C	D					
	Do baseline measures exist? If baseline data exists, Are the <b>main people responsible for the M&amp;E framework</b> available and prepared to support the evaluation?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Is the sampling process clear?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Are intended and actual beneficiaries distinguished in the baseline data?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Is there a record of who was involved in what project activities and when? ( <b>Roles and Responsibilities</b> )	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Are the data collection instruments available?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Is data (qualitative/quantitative) collected for all stipulated indicators? Any missing data? Are the <b>M&amp;E indicators aligned with the key components</b> set out in the Theory of Change/Logical framework?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	How frequently is the data collected and <b>analysed, and by whom and how?</b>	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Are the measures being used reliable i.e. is measurement error likely to be a problem?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	If baseline data is not yet available, are there specific plans for when baseline data would be collected and how feasible are these?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	Any relevant <b>national statistics</b> available?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
2.3.2 Is secondary data available?	Data from <b>previous evaluations</b> available done by others on same/similar projects?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
2.3.3 Is GESI/CS/ HRBA disaggregated data available?	In the baseline?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	For each of the indicators during project?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	
	In any mid-term or process review?	3 2 1	3 2 1	3 2 1	3 2 1		3 2 1	3 2 1	



Components		Evaluability Assessment Questions				Outcome (A to D Average)				Remarks on Assessment
Accessibility to and availability of stakeholders	Who are the stakeholders and how are they linked? E.g. linked financially, operationally or otherwise. Map them, identifying key decision makers.	3	3	3	3	3	3	3	3	
		2	2	2	2	2	2	2	2	
		1	1	1	1	1	1	1	1	
	Are there <b>physical security risks</b> ?	3	3	3	3	3	3	3	3	
Resources available to do the evaluation	Will <b>weather</b> be favourable?									
	Are they <b>motivated to get involved in the evaluation</b> or any likely resistance? Check stakeholder buy-in.	3	3	3	3	3	3	3	3	
		2	2	2	2	2	2	2	2	
		1	1	1	1	1	1	1	1	
Coordination requirements?	<b>Time available?</b>	3	3	3	3	3	3	3	3	
		2	2	2	2	2	2	2	2	
		1	1	1	1	1	1	1	1	
	<b>Funding available</b> for the relevant team and duration?	3	3	3	3	3	3	3	3	
Coordination requirements?	People with the <b>necessary skills</b> available at this point?	3	3	3	3	3	3	3	3	
		2	2	2	2	2	2	2	2	
		1	1	1	1	1	1	1	1	
	How many other donors, government departments, researchers or NGOs need to be or want to be involved in the evaluation (without compromising the need for reasonable degree of objectivity)? Are we clear?	3	3	3	3	3	3	3	3	
Coordination requirements?	What <b>forms of coordination</b> are possible and/or required?	3	3	3	3	3	3	3	3	
		2	2	2	2	2	2	2	2	
		1	1	1	1	1	1	1	1	
TOTAL(Section 3)										

Note: Blue coloured rows are the important questions to be addressed.

# Costing an Evaluation: Budget considerations and calculation for evaluations

A: Evaluation team Costs		# Days	Daily Rate	Total Cost
Professional fees	Team Leader/			
	Evaluator 1			
	Evaluator 2			
	Total			
Flights (International)	Evaluator 1			
	Evaluator 2			
	Total			
Per Diem costs (time in the field)	Evaluator 1			
	Evaluator 2			
	Total			
TOTAL A				

B: Evaluation Implementation and Data Collection costs		#	Cost	Total
Internal Flights				
Car Hire				
Translation Focus Group and workshop related costs				
Other costs				
TOTAL B				

C: Evaluation distribution Costs		#	Cost	Total
Report Editing				
Report Publication				
Stakeholder meeting				
TOTAL C				

Total Evaluation Costs (A+B+C)



# Terms of Reference:

[OPTIONAL DESCRIPTOR] EVALUATION OF THE  
[NAME OF PROGRAM/PROJECT OR COMPONENT TO BE EVALUATED]

## A. Background of Program/Project

This background should provide a summary of the program/project to be evaluated, its objectives, history, target beneficiaries, implementors, funding. The description should also contextualize the program/project in terms of its outputs and delivery mechanism, its outcomes and how these link to broader development goals (under the Philippine Development Plan and Sustainable Development Goals).

The program/project context, performance, and issues should also be discussed in order to provide an idea of the additional knowledge needed for program/project improvement. Such performance and issues may cite existing evaluations, studies, and administrative reports.

Finally, a background of the evaluation, including its purpose and which agency is commissioning the same and the agency’s stakeholders involved in the evaluation project, should be described in brief.

## B. Evaluation Objectives and Questions

The first paragraph should describe the main purpose(s) for the evaluation, particularly its policy mandate, intended use, and main subject of inquiry that is further fleshed out into the following evaluation questions.

Relevance	
Efficiency	
Effectiveness	
Equity	
Sustainability	
Others	

The TOR may note that the evaluation questions are indicative and may be prioritized or enhanced based on feasibility. It is recommended to have less than ten (10) evaluation questions in the TOR, prioritizing those which are of interest to stakeholders and those which are feasible given evaluability considerations. The consultant’s bids and inception report are expected to have more detailed questions. Refer to the Evaluation Plan in the annex, as needed.

## C. Scope of Services and Methodology

The first part should describe the broad type(s) and approach of the evaluation—i.e., if a formative, process, or summative/impact evaluation—and spell out the scope of the evaluation in terms of the time period of the program/project, depth of coverage including components, and target groups. It may also be useful to specify aspects which are outside the scope of the evaluation.

The next part should describe the overarching methodology to be employed and the data collection and analysis methods required. While the methodological framework should be specific enough in order to set realistic expectations, it should also be flexible enough to give bidders some room to propose additional or alternative proposals. It is recommended that minimums be defined, e.g., in terms of survey sample, number of case studies, key informants to be interviewed, stakeholders to be engaged, etc. The minimum activities to be undertaken by the firm should also be spelled out. Refer to the Evaluation Plan in the annex, as needed.

## D. Deliverables and Schedule of Expected Outputs

This section should spell out the overall duration of the contract and the outline the key milestone submissions—at the minimum, inception report, draft evaluation report, final evaluation report, and any interim outputs such as monthly progress reports—the description of the minimum contents of these submissions and format. Required reviews and approvals shall also be outlined, following the table below:

Deliverables/Outputs	Target Due Dates	Review and Approval

E. Governance and Accountability

This section should spell out the roles and responsibilities of the evaluation consultant and the evaluation management team—particularly the evaluation manager and evaluation commissioner. Any reference group or technical working group assembled to guide the evaluation project should also be described and their composition spelled out. Reporting and supervision lines and coordination mechanisms, including regular progress updating, should be spelled out. Finally, facilities, information, data, and other support to be provided by the evaluation manager to the evaluation consultant must be spelled out.

Approvals—e.g., ethics reviews and survey instrument clearances—must also be spelled out. It is also important to reiterate ethical norms and standards that are specified in the National Evaluation Policy Framework (NEPF) and its Guidelines.

F. Professional Qualifications of the Successful Contractor and its Key Personnel

1. Present the minimum qualifications of the firm or institution to be sought, including applicable experience in terms of years in business and previous engagements or contracts with respect to the project at hand. The previous engagements or contracts should be able to show the thematic specialization(s) of the firm as well as their capacity for evaluation.
2. Present the minimum composition of the evaluation team that the firm will assign to the project, their respective roles, and minimum qualifications in terms of academic achievement, professional background, and previous undertakings. It is recommended that the project lead or lead investigator is an incumbent or “organic” personnel of the firm.

G. Scope of Price Proposal and Schedule of Payments

1. Specify the Approved Budget for the Contract (ABC) of the project.
2. Specify the type of contract—e.g., lump sum price contract, time-based contract, etc.—noting that all consultancy contracts must be fixed-price output-based contracts regardless of extensions. The cost structure should be spelled out, i.e., if all costs will be incorporated in the remuneration, or if reimbursable costs will be allowed. Specify the minimum cost components.
3. In a table, present the payment terms in terms of deliverables or outputs to be submitted, the target date for submission and for approval, and equivalent proportion of the lump sum price. Specify also if an advanced payment for mobilization will be allowed.

Period	Deliverables/Outputs	Target Due	Percentage of Lump Sum Price

H. Criteria for Shortlisting and for Evaluation of Bids

1. Specify the process and criteria for shortlisting.

	Score
<b>Experience of the Firm</b> , particularly years of operation and past projects that are relevant	50
<b>Qualification of Key Staff to be Assigned</b> , ideally of the project lead and an incumbent employee of the firm	30
<b>Current Workload Relative to Capacity</b> , which may include a measure of financial capacity	20

2. For the Evaluation of Bids, Specify the evaluation scheme, i.e., if Quality-Based or Quality Cost-Based. With the latter, specify the distribution of the scores for the technical and financial proposals, with the latter taking a minimum of 15 percent and maximum of 40 percent. Specify that the minimum passing score of the technical proposal is 70 percent and spell out the criteria for the technical evaluation, as follows:

	Score		
	Sample 1: Balanced Rating	Sample 2: Emphasis on Methodology	Sample 3: Emphasis on Capacity
<b>I. Quality of Evaluation Team</b> Overall: composition of the team & allocation of tasks: Lead Evaluator: a. Experience b. Training c. Education  Evaluator 1 a. Experience b. Training c. Education  Evaluator 2 a. Experience b. Training c. Education	40	30	50
<b>II. Experience and Capability of Consultant (Firm)</b>  A. Capacity of the firm in terms of experience, organization and leadership B. Performance in previous projects, quality of outputs, and relationship with past clients C. Financial capacity	30	10	30

<b>III. Plan of Approach and Methodology</b>  A. Clarity and completeness of proposed approach and methodology to respond to the TOR requirements B. Innovativeness of proposed methodologies to respond to the evaluation study objectives/ questions C. Feasibility of the work plan given the approach and methodology and the risks to the project	30	60	20
--	----	----	----

3. In the combined scoring, the Financial Proposal will be computed as a ratio of the Proposal’s offer to the lowest price among the proposals received.

I. Annexes to the TOR

Append annexes to the TOR, which should include the following at the minimum

- 1. Annex 1. Evaluation Plan
- 2. Annex 2. Indicative timeline based on the schedule of outputs and deliverables
- 3. Others: project background documents, proposals, reports, etc.
- 4. Refer to the National Evaluation Policy Framework (NEPF) and the Guidelines

# Sections of an evaluation study TOR

Summary description	Detailed guidance
1. Background knowledge and rationale <ul style="list-style-type: none"><li>■ Describe program/project with key milestones in its history</li><li>■ Show context within broader development strategy at national/sector/local level</li></ul>	<ul style="list-style-type: none"><li>■ Current purpose, objectives, and intended outcomes key output, outcome, and impact indicators</li><li>■ Rationale for the evaluation and overarching evaluation objective and question, include overview of decisions likely influenced by findings</li><li>■ History of the program, include how objectives and targeted outcomes have changed over time</li><li>■ Context in which the program is situated, include organizational, social, political, regulatory, economic, or other factors that are relevant to the program’s implementation</li><li>■ Roles and responsibilities of various key stakeholders in designing and implementing the program, noting any significant changes that have occurred in these roles over time</li><li>■ Any studies or evaluations that have been conducted on the program or related activities to date – attach the M&amp;E framework for the program, if available</li></ul>
2. Objectives and specific evaluation questions <ul style="list-style-type: none"><li>■ Present information about why the evaluation is being conducted</li></ul>	<ul style="list-style-type: none"><li>■ Max of 3-5 objectives – simpler, shorter is better</li><li>■ Use outcome-focused language</li><li>■ Define specific questions for each</li></ul>

<ul style="list-style-type: none"><li>■ State objectives of the evaluation</li><li>■ Define specific evaluation questions</li><li>■ Describe expected users of the evaluation</li><li>■ national/sector/local level</li></ul>	<p>objective – maybe broad or specific depending on the objective</p> <ul style="list-style-type: none"><li>■ Keep list of questions to a minimum</li></ul>
3. Scope, approach, and methodology <ul style="list-style-type: none"><li>■ Specify the scope of the evaluation</li><li>■ Outline broad approach to set realistic expectations</li><li>■ Define degree of flexibility evaluator has in proposing additional or alternative methods</li></ul>	<ul style="list-style-type: none"><li>■ Define time period; depth of coverage; target groups; outside of scope</li><li>■ Overarching methodological framework, e.g., case study, sample survey, desk review, mixed methods</li><li>■ Expected data collection and analysis methods, with descriptions of any instruments used to collect needed information</li><li>■ Outcome and output indicators used to measure performance, along with associated baseline and target data</li><li>■ Availability of other relevant data, e.g., existing local, regional, or national data, or data from similar programs</li><li>■ Process for verifying findings with key stakeholders</li><li>■ Meetings expected with particular stakeholder groups</li><li>■ Define how various users/ stakeholders in the evaluation are likely to be involved</li></ul>

Summary description	Detailed guidance
4. Governance and accountability arrangements <ul style="list-style-type: none"><li>■ Outline roles and responsibilities for the evaluation and the management and coordination arrangementsmethods</li></ul>	<ul style="list-style-type: none"><li>■ Describe organization and function of decision-making structures, e.g., technical working group, reference group; participation of other stakeholders (e.g., beneficiary representatives to validate results); and lines of accountability with clear guidance on who will review and approve the implementation plan for the evaluation and other outputs of the study</li><li>■ Mention any specific support or resources being provided by NEDA-CO or RO, e.g., cover letters for data collection site-visit coordination with program implementation agencies.</li></ul>
5. Guiding principles or values <ul style="list-style-type: none"><li>■ Specify research ethics or procedures that the evaluator is expected to follow</li></ul>	
6. Deliverables, schedule and budget	<ul style="list-style-type: none"><li>■ Specify expected deliverables, timeline, and work plan – may request the evaluator to provide a detailed timeline and milestones within the timeline specified</li><li>■ List outputs that should be delivered</li><li>■ Include details related to format, content, length, intended audience, and the expected review process</li><li>■ State the budget available for the evaluation and what the budget covers</li></ul>

	<ul style="list-style-type: none"><li>■ Define type of contract (e.g., time-based, lump sum, output-based)</li></ul>
7. Qualifications of the evaluator	<ul style="list-style-type: none"><li>■ Present the expected professional profile of the evaluation team, including desired experience and credentials, noting the minimum professional requirements or competencies</li><li>■ Refer to relevant sections of GPRA for any other requirements</li></ul>

# Inception Report Template

I. Program/Project Information	
1. Name of Program/Project	(if thematic evaluation, identify the theme and the component programs, if any, that will be evaluated)
2. Program/Project Location(s)	
3. Program/Project Duration	
4. Lead Government Agency	
5. Other Government Agencies	
6. Other Implementing Partners	
7. Alignment with the Philippine Development Goals (identify PDP Chapters relating to the program)	
8. Alignment with the Sustainable Development Goals (identify SDGs relating to the program)	

II. Background and context
Describe the nature, scale, and context of the intervention that is being evaluated, including the critical social, economic, political, geographic and demographic factors that influenced its design and implementation.

III. Evaluation bjective, purpose and scope
Clearly state the objectives of the evaluation, potential users of the evaluation findings and recommendations, including justification for the timing and usefulness of the evaluation

IV. Evaluation criteria and questions
Outline the key evaluation questions grouped according to the OECD DAC criteria of i) relevance, ii) effectiveness, iii) efficiency, iv) sustainability, and v) impact

V. Evaluation approach and methodology
Describe the overall methodological approach for conducting the evaluation, including data sources, sampling strategy, and data collection tools.

VI. Evaluation matrix
Tool for evaluators as basis for planning and conducting an evaluation. It details evaluation questions per evaluation criteria, data sources, success indicators, as well as data collection and data analysis methods

Evaluation criteria	Key questions	Sub questions	Data sources	Data collection methods	Success indicators	Data analysis methods

VII. Schedule of key milestones
Include a table or Gantt chart listing all tasks and deliverables related to the conduct of the evaluation, including activities that involve the evaluation manager, commissioner, and ERG. Activities can be grouped according to the following phases: <div><div>1. Desk review and inception report</div><div>2. Data collection mission</div><div>3. Evaluation report writing</div><div>4. Dissemination and communication</div></div>



Activity	Estimated no. of days	Date of completion	Data sources

# Evaluation Report Template<sup>1</sup>

[The evaluation report is the key product of the evaluation process. Its purpose is to provide a transparent basis for accountability for results, for decision-making on policies and programmes, for learning, for drawing lessors learnt and for improvement.]

## 1. Basic Details

Evaluation Title:			
Evaluation Type:			
Project/Programme Title:			
Project Start Date:		Project End Date:	
Geographical Coverage:	[Indicate if coverage covers multiple Regions, Provinces, etc.]		
Executing Agencies:			
Evaluator(s):			
Evaluation Duration (months) :	[Indicate duration of the evaluation.]		
Commissioning Unit:	[e.g. NEDA-CO (SDS, MES,PIS,etc.) or NEDA-RO XI]		
NEDA Focal Point:		Contact Details:	
Submission Date:			

<sup>1</sup> Drawn from multiple sources: UNEG, Norms and Standards for Evaluation, 2016; World Bank, Managing Evaluations: A How-To Guide for Managers and Commissioners of Evaluation, 2015; ADB Guidelines for the Preparation of Country Assistance Program Evaluations and Country Partnership Strategy Final Review and Validations, 2015; UNODC, Evaluation Handbook, 2012; UNDP, Handbook on Planning, Monitoring, and Evaluating for Development Results, 2009; Austrian Development Cooperation, Guidelines for Project and Programme Evaluation, 2009

2. Table of Contents

[Should always include boxes, figures, tables and annex with page references]

3. Acknowledgements [max 500 words]

|

|

**4. List of Acronyms and Abbreviations**

**5. Executive Summary** [max 750 words]

[Section should consist of a concise executive summary that includes: 1. An introduction and the background: short description of the project/programme evaluated including its objectives; 2. The major findings of the evaluation; 3. The main conclusions; 4. The major recommendations; 5. The major lessons learned]

**6. Introduction** [max 750 words]

[1. Discuss the overall concept and design of the project/programme. 2. Identify the primary audience or users of the evaluation, identify the interventions (projects/programmes/policies) that were evaluated and are relevant for the current evaluation. 3. Acquaint the reader with the structure and contents of the report and how the information contained in the report will meet the purpose of the evaluation and satisfy the information needs of the report’s intended users.]

**7. Description of the Program** [max 750 words]

[1. Describe what is being evaluated, who seeks to benefit, and the problem of issue it seeks to address. 2. Present the theory of change, results framework, and external factors likely to affect success. 3. Link the intervention to national priorities, identify the phase in implementation of the intervention and any significant changes that have occurred over time, and explain the implications of those changes for the evaluation. 4. Identify the key partners involved in the implementation. 5. Describe the scale of the intervention and the size of the target population for each component. 6. Indicate total resources (human and financial). 7. Describe the context of social, political, economic, and institutional factors, and the geographical landscape within the intervention operates and explains the effects those factors present for its implementation and outcomes. 8. Point out design weaknesses or other implementation constraints.]

**8. Purpose of the Evaluation** [max 350 words]

[Explain why the evaluation was conducted, why the intervention is being evaluated at this point in time, and why it addressed questions it did.]

**9. Key Questions, Scope and Limitations** [max 700 words]

[1. Indicate the key evaluations questions that were used to assess the overall performance of the project/programme. 2. Define the parameters of the evaluation (e.g. time period, segments of targeted population) and which components, outputs, or outcomes were and were not assessed. 3. Define the evaluation criteria or performance standards used and the rationale for selecting the particular criteria used in the evaluation.]

**10. Methodology** [max 750 words]

[1. Provide a statement of methods used to obtain, collect, and analyze the data; 2. Discuss the standard of measure that will be used to evaluate the performance relative to the evaluation question<sup>2</sup>; 3. Discuss the level of stakeholder engagement in the evaluation and how the level of involvement contributed to the credibility of the evaluation and the results; 4. Discuss the measures taken to protect the rights and confidentiality of informants; 5. Discuss the composition of the evaluation team, the background and skills of team members and the appropriateness of the technical skill mix, gender balance and geographical representation for the evaluation; 6. Discuss the major limitations of the methodology should be identified and openly discussed as to their implications for the evaluation as well as steps to mitigate those limitations. Note: Any tools used should be put to the annex and be referenced accordingly.]

**11. Findings**

[Indicate the analysis of information and articulate the all relevant findings of the evaluation. Should be presented as statements of fact that are based on analysis of the data. They should be structured around the evaluation criteria and questions developed. Variances between planned and actual results should be explained as well as factors affecting the achievement of intended results. Assumptions or risks in the project or programme design that subsequently affected implementations should also be discussed.]

**12. Summary and Explanation of Findings and Interpretations**

[max 700 words]

**13. Conclusions and Recommendations** [max 750 words]

[1. Conclusions should be comprehensive and balanced, and highlight the strengths, weaknesses and outcomes of the intervention. Moreover, they must be able to respond to key evaluation questions and provide insights into the identification or solutions to important problems or issues pertinent to the decision making of intended users. 2. Recommendations should emerge logically from the evaluation’s findings and conclusions. Moreover they should be relevant, realistic, with clear priorities for action.]

**14. Lessons Learned, Generalizations, and Alternatives** [max 750 words]

[A forward looking discussion of new knowledge gained from the particular circumstance (intervention, context outcomes, even about evaluation methods) that are applicable to a similar context. ]

**15. Annexes**

[Annexes should include: 1. TOR for the evaluation; 2. Additional methodology-related documentation, such as the evaluation matrix and data collection instruments (questionnaires, interview guides, observation protocols, etc.) as appropriate; 3. List of individuals or groups interviewed or consulted and sites visited. 4. List of supporting documents reviewed; 5. Project/Programme results map or results framework; 6. Short biographies of the evaluators and justification of team composition; 7. Code of conduct signed by evaluators.]

<sup>2</sup> A summary matrix displaying for each of evaluation questions, the data sources, the data collection tools or methods for each data source and the standard measure by which each question was evaluated should be developed and included in the annex

# Evaluation Report Checklist

Checklist for Evaluation Reports		
Program/Project:		
Agency:		
Evaluator:		
1. The Report Structure		
1.0	The Report is well structured, logical, clear and complete.	<input type="checkbox"/>
1.1	Report is logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations).	<input type="checkbox"/>
1.2	The title page and opening pages provide key basic information:	
	<input type="checkbox"/> name of the evaluation	<input type="checkbox"/>
	<input type="checkbox"/> timeframe of the evaluation and date of the report	<input type="checkbox"/>
	<input type="checkbox"/> name of evaluator(s)	<input type="checkbox"/>
	<input type="checkbox"/> name of the agency commissioning the evaluation	<input type="checkbox"/>
	<input type="checkbox"/> table of contents which also lists tables, graphs, figures and appendices	<input type="checkbox"/>
	<input type="checkbox"/> list of terminology including acronyms	<input type="checkbox"/>
1.3	The Executive Summary is a stand-alone section of 2-3 pages that includes:	
	<input type="checkbox"/> overview of the evaluation	<input type="checkbox"/>
	<input type="checkbox"/> evaluation objectives and intended audience	<input type="checkbox"/>
	<input type="checkbox"/> evaluation methodology	<input type="checkbox"/>
	<input type="checkbox"/> most important findings and conclusions	<input type="checkbox"/>
	<input type="checkbox"/> main recommendations	<input type="checkbox"/>
1.4	Appendices increase the credibility of the evaluation report. Appendices may include:	
	<input type="checkbox"/> evaluation terms of reference	<input type="checkbox"/>
	<input type="checkbox"/> list of persons interviewed and sites visited	<input type="checkbox"/>
	<input type="checkbox"/> list of documents	<input type="checkbox"/>
	<input type="checkbox"/> further information on the methodology, such as data collection instruments, including details of their reliability and validity	<input type="checkbox"/>
	<input type="checkbox"/> evaluator's justification of team composition	<input type="checkbox"/>

2. Full description of Program		
2.0	The report presents a clear and full description of the evaluated program.	<input type="checkbox"/>
2.1	The inputs, outputs and results of the program are clearly described based on the Program Logic Map.	<input type="checkbox"/>
2.2	The context of key social, political, economic, demographic, and institutional factors that have a direct bearing on the evaluated program is described.	
2.3	The scale and complexity of the evaluated program are clearly described. For example:  <input type="checkbox"/> the number of components, if more than one, and the size of the population each component is intended to serve, either directly or indirectly.  <input type="checkbox"/> the geographic context and boundaries (such as the region, and/or landscape and challenges where relevant).  <input type="checkbox"/> the purpose and goal, and organization/management of the program and its parts.  <input type="checkbox"/> total resources from all sources, including human resources and budget(s) including Agency, State and Commonwealth funding.	<input type="checkbox"/>
2.4	The key stakeholders involved in the program, including the implementing agency(s) and partners, other key stakeholders and their roles.	<input type="checkbox"/>
2.5	The report identifies the implementation status of the program, including its phase of implementation and any significant changes (e.g. plans, strategies) that have occurred over time. In addition, explain the implications of those changes for the evaluation.	<input type="checkbox"/>
3. Evaluation Purpose and Scope		
3.0	The evaluation's purpose, objectives and scope are fully explained.	<input type="checkbox"/>
3.1	The purpose of the evaluation is clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed and how the information will be used.	<input type="checkbox"/>
3.2	The report should provide a clear explanation of the evaluation's 'SMART' results and scope including key evaluation questions. It should describe and justify what the evaluation did and did not cover.	<input type="checkbox"/>

4. Evaluation Process		
4.0	The report describes the evaluation process and clearly explains how the evaluation was designed to address the results criteria and answer the evaluation questions.	
4.1	The report describes the data collection methods and analysis, the rationale for selecting them, and their limitations. Baseline data and benchmarks are included where relevant.	<input type="checkbox"/>
4.2	The report describes the data sources, the rationale for their selection, and their limitations. It includes a discussion of how a mix of data sources was used to obtain a diversity of perspectives, ensure data accuracy, validity and overcome data limitations.	<input type="checkbox"/>
4.3	The report gives a complete description of the stakeholder consultation process during the evaluation, including the rationale for selecting the particular level of consultation.	<input type="checkbox"/>
5. Findings		
5.0	Findings relate directly to the results criteria.	<input type="checkbox"/>
5.1	Reported findings reflect systematic and appropriate analysis and interpretation of the data.	<input type="checkbox"/>
5.2	Reported findings address the 'SMART' result criteria (such as efficiency, effectiveness, sustainability, impact and relevance) and key questions defined in the evaluation scope.	<input type="checkbox"/>
5.3	Findings are objectively reported based on the evidence.	<input type="checkbox"/>
5.4	Gaps and limitations in the data and/or unanticipated findings are reported and discussed.	<input type="checkbox"/>
5.5	Overall findings are presented with clarity, logic, and coherence.	<input type="checkbox"/>

6. Conclusions		
6.0	Conclusions present reasonable judgments based on findings and substantiated by evidence.	
6.1	The conclusions reflect reasonable evaluative judgments relating to key evaluation questions.	<input type="checkbox"/>
6.2	The conclusions provide insights into the identification and/or solutions to important problems or issues.	<input type="checkbox"/>
6.3	Conclusions present strengths and weaknesses of the program being evaluated, based on the evidence presented and taking due account of the views of a variety of stakeholders.	<input type="checkbox"/>
7. Recommendations		
7.0	Recommendations are relevant to the program and the evaluation's purpose, are supported by evidence and conclusions, and were developed with the involvement of relevant stakeholders.	<input type="checkbox"/>
7.1	The report describes the process involved in developing the recommendations including consultation with stakeholders.	<input type="checkbox"/>
7.2	Recommendations are based on evidence and conclusions.	<input type="checkbox"/>
7.3	Recommendations are actionable and reflect an understanding of the agency.	<input type="checkbox"/>
7.4	An implementation plan for the recommendations is included within the report.	<input type="checkbox"/>



Management response and implementation plan<sup>1</sup>

Program/Project Title:

Implementing Agency(ies):

Completion Date of Evaluation:

Date of Issuance of Management Response:

Contributors: Firm/s commissioned to conduct the evaluation

Cleared by: The Commissioning Agency/Unit

I. Context, background, and key findings

(This part may be a three- to four-paragraph summary that establishes the bases for recommendations and action points.)

II. Recommendations and management response

Evaluation Recommendation 1				
Management Response:				
Key action(s) to im- plement	Due date of implementation (specify target month, year)	Unit(s) Responsible (indicate their roles in imple- mentation)	Tracking	
			Status (i.e., not start- ed; ongoing; delayed; completed	Remarks (e.g., issues encountered in implement- ing the action points)
1.1				
1.2				
1.3				

Evaluation Recommendation 2				
Management Response:				
Key action(s) to im- plement	Due date of implementation (specify target month, year)	Unit(s) Responsible (indicate their roles in imple- mentation)	Tracking	
			Status (i.e., not start- ed; ongoing; delayed; completed	Remarks (e.g., issues encountered in implement- ing the action points)
2.1				
2.2				
2.3				

Evaluation Recommendation 3.				
Management Response:				
Key action(s) to im- plement	Due date of implementation (specify target month, year)	Unit(s) Responsible (indicate their roles in imple- mentation)	Tracking	
			Status (i.e., not start- ed; ongoing; delayed; completed	Remarks (e.g., issues encountered in implement- ing the action points)
3.1				
3.2				
3.3				

<sup>1</sup> Adopted from the UNDP Management Response Template

# Quality Assurance Checklist

Evaluation Quality Assessment Detail Report Checklist			
Scoring Sheet			
Code		Rubric for Assigning Rating for Each Criteria	Value
HS	Highly Satisfactory	All parameters were fully met and there were no shortcomings in the evaluation report	6
S	Satisfactory	All parameters were partially met with some shortcomings in the evaluation report	5
MS	Moderately Satisfactory	The parameters were partially met with some shortcomings in the evaluation report	4
MU	Moderately Unsatisfactory	More than one parameter was unmet with significant shortcomings in the evaluation report	3
U	Unsatisfactory	Most parameters were not met and there were major shortcomings in the evaluation report	2
HU	Highly Unsatisfactory	None of the parameters were met and there were severe shortcomings in the evaluation report	1
N/A	Not Applicable	Does not apply since the answer is not available	0

Basic Information on the Evaluation	
Basic Information	
Title of Evaluation Report:	
Type of Evaluation	
Commissioning unit:	
Philippine Development Plan Output/Outcome:	
Author/s of the evaluation report: (Name, Title, Email, Gender, Nationality)	
Evaluation budget proposed:	
Date of Completion:	
Number of pages (excluding annexes):	
Review Information	
Final quality assessment submission date and signature of the reviewer:	
Evaluation Manager approval date and signature:	

Summary of Quality Assessment Report	
Sections	Summary of comments
<b>1. Evaluation Structure and Design:</b> <i>Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria and key questions for the evaluation and give adequate time and resources?</i>	
<b>2. Evaluation Report &amp; Methodology:</b> <i>Are the evaluation reports objectives, criteria, methodology and data sources, fully described and are they appropriate given the subject being evaluated and the reasons for carrying out the study?</i>	
<b>3. Evaluation Results, Findings, Conclusions &amp; Recommendations:</b> <i>Does the report clearly and concisely outline and support its findings, conclusions and recommendations?</i>	

Terms of Reference Section Rating	
Terms of Reference (TOR) and Design	
Rating	Comments / Suggestions for Improvement
1. Does the TOR clearly outline the focus for the evaluation in a logical and realistic manner?	
2. Does the TOR detail timescales and budgets for the evaluation?	
3. Does the TOR clearly outline the evaluation’s planned approach?	
4. Is the proposed outline of the evaluation approach and methodology clearly detailed in the ToR?	

Report And Methodology Section Rating	
Report and Methodology	
Rating	Comments / Suggestions for Improvement
<b>1. Structure</b>	
<b>1.1 Is the evaluation report well-balanced and structured?</b> The report is logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations).	
<b>1.2 The Executive Summary is a stand-alone section of 2-3 pages that includes:</b> <ul style="list-style-type: none"> <li>a. Overview of the evaluation object</li> <li>b. Evaluation objectives and intended audience</li> <li>c. Evaluation methodology</li> <li>d. Most important findings and conclusions</li> <li>e. Main recommendations</li> </ul>	
<b>1.3 Annexes increase the credibility of the evaluation report.</b> They may include, inter alia: <ul style="list-style-type: none"> <li>a. TORs</li> <li>b. List of persons interviewed and sites visited.</li> <li>c. List of documents consulted</li> <li>d. More details on the methodology, such as data collection instruments, including details of their reliability and validity</li> <li>e. Evaluators biodata and/or justification of team composition</li> <li>f. Evaluation matrix</li> <li>g. Results framework</li> </ul>	
<b>2. Evaluation object</b>	
<b>2.1 The report presents a clear and full description of the ‘object’ of the evaluation.</b> The “object” of the evaluation is the intervention (outcome, programme, project, group of projects, themes, soft assistance) that is (are) the focus of the evaluation and evaluation results presented in the report.	

2.2	The logic model and/or the expected results chain of the object is clearly described.
2.3	<b>The scale and complexity of the object of the evaluation are clearly described, for example:</b> <ul style="list-style-type: none"><li>a. <b>The number of components</b>, if more than one, and the size of the population each component is intended to serve, either directly and indirectly.</li><li>b. <b>The geographic context and boundaries</b> (such as the region, country, and/or landscape)</li><li>c. The purpose and goal, and organization/management of the object</li><li>d. The <b>total resources</b> from all sources</li></ul>
2.4	The key stakeholders involved in the object implementation, including the implementing agency(s) and partners
3.	<b>Evaluation objective</b>
3.1	The purpose of the evaluation is clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, how the information will be used.
3.2	The report should provide a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover.
4.	<b>Methodology</b>
4.1	<b>Is the evaluation's methodological approach clearly outlined?</b> <i>The report presents transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve evaluation purposes.</i>

4.2	Is the evaluation methodology adequate for scope of the evaluation?
4.3	The report describes the sampling frame – area and population to be represented, rationale for selection, mechanics of selection, numbers selected out of potential subjects, and limitations of the sample.
4.4	The report presents evidence that adequate measures were taken to ensure data quality, including evidence supporting the reliability and validity of data collection tools (e.g. interview protocols, observation tools, etc.)
5.	<b>Data Collection</b>
5.1	Are data collection methods and analysis clearly outlined?
5.2	Is the data collection approach and analysis adequate for scope of the evaluation?
5.3	Are any changes to the evaluation approach or limitations in implementation during the evaluation mission clearly outlined and explained?

Recommendations and Lessons Section Rating	
Recommendations and Lessons Section	
Rating	Comments / Suggestions for Improvement
6.	<b>Findings</b>
6.1	Findings respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report and are based on evidence derived from data collection and analysis methods described in the methodology section of the report.

6.2	Reported findings reflect systematic and appropriate analysis and interpretation of the evidence
6.3	Reasons for accomplishments and failures, especially continuing constraints, were identified as much as possible
6.4	Does the evaluation include an assessment of the projects M&E design, implementation and overall quality?
6.5	Are all indicators in the logical framework / theory of change assessed individually, with final achievements noted?
7.	<i>Conclusions</i>
7.1	Conclusions present reasonable judgments based on findings and substantiated by evidence, and provide insights pertinent to the object and purpose of the evaluation.
7.2	Conclusions are well substantiated by the evidence presented and are logically connected to evaluation findings.
7.3	Conclusions present strengths and weaknesses of the object (policy, programmes, project's or other intervention) being evaluated, based on the evidence presented and taking due account of the views of a diverse cross-section of stakeholders.

8.	<i>Recommendations</i>
8.1	Recommendations are relevant to the object and purposes of the evaluation, are supported by evidence and conclusions, and were developed with the involvement of relevant stakeholders.
8.2	Are the recommendations clear, concise, realistic and actionable? Recommendations clearly identify the target group for each recommendation.
Lessons Learned	
What key lessons can be drawn from the report?	

Published with Support from



## **GUIDELINES ON EVALUATION IN THE NATIONAL GOVERNMENT**

### **Publish by**

#### **National Economic and Development Authority**

12 St. Josemaria Escriva Drive, Ortigas Center,  
Pasig City 1605 Philippines  
Tel: (+632) 8631-09-45 to 55  
Email: [info@neda.gov.ph](mailto:info@neda.gov.ph)  
[www.neda.gov.ph](http://www.neda.gov.ph)

#### **Department of Budget and Management**

General Solano Street., San Miguel, Malacañang,  
Manila Philippines  
Tel: (+632) 8657-3300  
Email: [osec@dbm.gov](mailto:osec@dbm.gov)