



Evaluation Operational Policy

For Official Development Assistance funded initiatives as part of New Zealand's International Development Cooperation

Updated April 2025

Contents

Introduction	2
Summary of key operational policy statements	2
Overview of key roles and responsibilities	3
What is evaluation?	4
Why evaluate?	5
What is evaluated?	5
Strategic evaluation	5
Activity evaluation	6
Evaluation quality	7
Dissemination	7
Publication of evaluation reports	7
Management response	8
Learning	8
Policy review	8
Appendix A: Evaluation prioritisation	9

Introduction

New Zealand contributes Official Development Assistance (ODA) to sustainable development and humanitarian action in Pacific Island and developing countries.¹ The majority of our ODA focuses on countries most in need, particularly small island developing states, and least developed countries. New Zealand's ODA is primarily managed by the New Zealand Ministry of Foreign Affairs and Trade (MFAT).

This Evaluation Operational Policy² sets out the roles, responsibilities, expectations, and standards for evaluations funded by New Zealand's ODA. It responds to recent changes such as programme governance and planning systems, the revised internationally agreed OECD-DAC criteria for evaluating development assistance, and New Zealand's International Cooperation for Effective Sustainable Development (ICESD) policy.

Summary of key operational policy statements

1. Evaluation is designed to assess whether development interventions are *relevant* to context, *coherent* with other interventions, *achieve* the objectives, *contribute* to outcomes sought, deliver results in an *efficient* way, and have positive impacts that *last*.
2. Evaluation enhances accountability and transparency and will inform continuous improvement. Evaluation is distinct from our monitoring processes.
3. Evaluation as it relates to New Zealand's ODA occurs at two levels:
 - **Strategic evaluation** aimed at programme level (PDG plan and thematic areas). These evaluations are funded by the Strategic Evaluation and Research Fund (SERF), which has a sub-allocation of 1% of the total International Development Cooperation (IDC) allocation for the 2024-2027 triennium. They are managed by the Monitoring, Evaluation, Research, and Learning (MERL) Unit. MFAT's Pacific and Development Leadership Team Strategy & Policy Group (PDLT S&P) provides oversight and sets the priorities for a rolling work plan, based on considerations of programme size, risk, demand for evidence and opportunities to improve.
 - **Activity evaluation** focused on activities or groups of activities within PDG Plans, primarily funded by the respective activity. These are managed by the Activity Manager under direction of the respective governance group, with support from the MERL Unit. Evaluations can be conducted at any point in the activity life cycle, depending on the purpose. Governance groups determine which activities in their PDG Plans should be evaluated based on prioritisation criteria in this policy.
4. Evaluation reports are considered complete once a management response has been endorsed by the respective PDG Plan owners. The report with the management response should be published within three months of completion. Any decision not to publish an evaluation report will be by exception, and only with PDLT S&P approval.³
5. MFAT's approach to evaluation is *utilisation-focused*, with the emphasis of evaluation work based on usefulness to intended users. The MERL Unit offers support to PDG

¹ ODA is defined by the OECD Development Assistance Committee (DAC) as government funding designed to promote the economic development and welfare of developing countries. New Zealand's support under Vote ODA is more expansive than the OECD DAC definition, as developed Pacific Island countries are also eligible.

² This policy replaces the previous Evaluation Policy, 2014.

³ All evaluation reports are subject to the Official Information Act 1982, even if not published.

Plan and thematic teams including independent internal technical advice, capacity building, and review of evaluation quality and processes.

Overview of key roles and responsibilities

The table below summarises key functions, roles and responsibilities for strategic and activity evaluations associated with New Zealand's ODA.

FOCUS	STRATEGIC EVALUATIONS	ACTIVITY EVALUATIONS
PRIORITISATION AND PLANNING		
Who decides and how?	PDLT S&P sets the priorities for a rolling work plan of PDG Plan and thematic evaluations, based on considerations of programme size, risk, demand for evidence and continuous improvement.	Activity managers (with support from Unit Managers and MERL) plan for evaluations during the design phase of their activities, which is then approved at the Business Case stage.
Funding	SERF	Activity/programme budget
GOVERNANCE AND OVERSIGHT		
Senior management oversight	MERL Unit reports to PDLT S&P on the progress against the strategic evaluation programme.	Activity managers record planned evaluations in Enquire. MERL Unit submits report six-monthly to PDG Governance Group on trends and progress of activity evaluations across PDG.
Management response	Unit Managers (MERL and relevant PDG Plan/programme) review and endorse for approval by Divisional Managers (DCI and relevant PDG Plan/programme) prior to publication.	Reviewed and endorsed by the Unit Manager of the activity for approval by the Divisional manager (PDG Plan owner).
IMPLEMENTATION OF EACH EVALUATION		
Who is accountable	Divisional Managers (DCI and the PDG Plan/programme)	Divisional manager of the activity.
Who is responsible	Unit Managers (MERL and the PDG Plan/programme)	Unit manager of the activity.
Who manages	MERL Advisers	Activity managers (with technical support from MERL).
Who conducts	Typically, evaluations are conducted by independent evaluation teams contracted externally from MFAT's MERL Panel and List of Suppliers. Sometimes, evaluations can be done internally by MFAT's MERL Unit, or jointly with external evaluators.	Typically, evaluations are conducted by independent evaluation teams contracted externally from MFAT's MERL Panel and List of Suppliers. Sometimes, evaluations can be done internally by MFAT's MERL Unit, or jointly with external evaluators. The MERL Unit can also support other teams to undertake self-evaluation.

FOCUS	STRATEGIC EVALUATIONS	ACTIVITY EVALUATIONS
DISSEMINATION AND USE		
Publication of evaluation report	MERL publishes on MFAT website and records in Enquire.	Activity manager informs MERL Unit, who then publishes on MFAT website. Activity manager records in Enquire.
Report back on findings, recommendations, response and follow-up	Evaluators, in consultation with MERL, prepare communication and dissemination plan. MERL reports to PDLT S&P and relevant Governance Group. The relevant Governance Group and PDLT S&P oversee follow-up actions required.	Evaluators, in consultation with Activity managers prepare communication and dissemination plan (with support from MERL). MERL Unit submits report six-monthly to PDG Governance Group on trends and progress of activity evaluations across PDG.
Use of findings in business cases and design	Governance Groups assess use of evaluation evidence in business cases. MERL commissions/produces regular syntheses of results and lessons from activity monitoring and completion assessments (AMAs and ACAs), and evaluations. Enquire report available with links to all evaluation reports searchable by PDG Plan, country, theme, and sector.	

What is evaluation?

MFAT uses a definition of evaluation adapted from the OECD-DAC evaluation criteria:

Evaluation is an evidence-based, systematic and objective assessment of an ongoing or completed project, programme, strategy or policy, its design, implementation and results, in relation to specified evaluation criteria.

Evaluation is designed to assess whether interventions are *relevant* to context, *coherent* with other interventions, *achieve* the objectives, *contribute* to outcomes sought, deliver results in an *efficient* way and have positive impacts that *last*. This aligns to internationally agreed criteria for evaluating development assistance and the development principles in New Zealand's ICESD policy (see following section).

Evaluations are an in-depth examination that takes place on a periodic basis. They are distinct from, but draw on, monitoring, which is a continuous process of assessing whether an investment or programme is on track to achieve intended outcomes.

Evaluations can be conducted at any point depending on the purpose. For example:

- **Formative** evaluations are conducted early or mid-way to inform learning and improvements to delivery.
- **Summative** evaluations are conducted at the conclusion and focus more on accountability.
- **Impact** evaluations could be conducted several years after completion to assess impact on longer term outcomes.

Why evaluate?

Evaluations are undertaken to:

- support evidence-informed design and delivery of our development assistance by learning about what works, what does not, and why⁴;
- support continuous improvement and adaptation to maximise value;
- promote accountability and transparency to Parliament, New Zealanders, our partner countries and other stakeholders for the resources used and the difference they have made – this is also important for New Zealand’s reputation as a member of the international aid community;
- develop an evidence base for subsequent activities or activity phases.

The use of evaluation evidence therefore focuses on continuous improvement, accountability and transparency. These are essential to the four development principles that underpin MFAT’s ICESD policy statement: Effectiveness, Inclusiveness, Resilience and Sustainability.

What is evaluated?

Evaluation as it relates to New Zealand’s ODA occurs at two levels.

Strategic evaluation

PDG Plans are plans that articulate programmes of work to progress outcomes set out in individual country strategies. PDG Plans relate to bilateral, multi-country or regional engagement. **Thematic programmes** refer to work not limited to an individual PDG Plan and often cover multiple countries.

A strategic independent evaluation of each PDG Plan and thematic programme will be conducted on a rolling cycle to assess progress against the intended outcomes, challenges and lessons learned.

PDLT S&P will oversee the strategic evaluation programme, and sets the priorities within the rolling cycle. The MERL Unit manages each strategic evaluation and reports back to PDLT S&P on progress against the evaluation programme.

Management responses for each strategic evaluation are provided to PDLT S&P for noting.

Responsibility for each strategic evaluation is shared jointly by the MERL Unit Manager and the relevant PDG Plan/thematic programme owner. Strategic evaluations are funded from the SERF sub-allocation.

⁴ This is an important part of the Performance Improvement Framework promoted by the Public Services Commission and is a priority for the Ministry of Foreign Relations and Trade.

Activity evaluation

Activities are specific investments within each PDG Plan and are designed to produce the results required to achieve outcomes as set out in the PDG Plan. Activities have their own Monitoring, Evaluation, Research and Learning (MERL) frameworks and are subject to annual monitoring of progress and reporting on completion. These activity monitoring and completion assessment (AMA and ACA) reports are an important evidence source that feeds into evaluations.

Decisions regarding which activities should be evaluated and when are made by the Activity managers (with support from relevant Unit Manager and MERL). Their decision is based on a set of criteria reflecting the activity's risk, complexity, length of implementation, level of investment, and strategic value. These considerations should be made during activity business case assessment, including the use of findings and recommendations arising from relevant prior monitoring reports.

The type of evaluation will be determined by the purpose. This ranges from fully external independent evaluation, internal evaluation, rapid reviews, and/or self-evaluation. Evaluation can be conducted at any stage in the activity life-cycle in order to inform learning, decision-making and accountability.

Activity managers are responsible for managing the evaluation and submitting the final report to their relevant Unit Managers and Divisional Manager. This includes preparing a management response to the evaluation recommendations and publishing it on the Ministry's website along with the evaluation report.

Funding for evaluations at the activity level generally comes from within the activity budget.

Evaluation quality

To ensure quality and use, evaluations are guided by the following core [principles](#) and criteria set out by the OECD-DAC and New Zealand's ICESD policy statement:

CORE PRINCIPLES	EVALUATION CRITERIA
<ul style="list-style-type: none"> • Clarity of intent • Credibility and rigour • Utility • Culturally and contextually responsive • Partnership and participation • Inclusion • Resilience • Donor cooperation and collaboration • Safeguarding and ethical practice • Strengths-based approach to capacity building 	<ul style="list-style-type: none"> • Relevance • Coherence • Effectiveness • Efficiency • Sustainability • Impact <p>Note: not all criteria will be relevant to every evaluation</p>

The MERL Unit supports PDG teams in a range of ways:

- **Independent internal technical advice** in the commissioning and design of evaluations. MERL Advisers work collaboratively with external evaluators and PDG Plan managers across all stages of the evaluation, including providing guidance on the level and type of evaluation, the terms of reference to conduct an evaluation, through to reviewing of draft and final reports, and management responses.
- **Capacity building** through evaluation training and preparing of practical evaluation resources to equip staff (both in Wellington and at Post) and our implementation partners with the knowledge and tools needed to plan, procure, and manage high quality evaluations.
- **Review of evaluation technical quality and process** to assess robustness of evaluation analysis, findings, and recommendations, and identify lessons to continuously improve the procurement, commissioning, implementation and reporting of evaluations.

Dissemination

Publication of evaluation reports

Transparency is central to effective development practice. In keeping with this principle, all **strategic evaluation** reports are published, in full or in summary form.

Activity evaluations suitable for publication are also published in full or in summary form. Any decision taken not to publish an evaluation will be by exception and put to the MERL Unit Manager for endorsement, and to DCI Divisional Manager for approval.

Publication must be on the Ministry of Foreign Affairs and Trade external website and within three months of completion, inclusive of the management response.⁵ The extent to which completed evaluations are published within the three-month timeframe is one of the Ministry's organisational performance indicators, reported in the Ministry's annual report to Parliament.

Management response

A management response is required for all evaluations. The response articulates the Ministry's assessment of the key findings, conclusions, and recommendations, and identifies what action will be taken as a result, and timeframes for these actions. The response is published alongside the evaluation report.

The findings, recommendations, and management response for Strategic Evaluations are presented to the relevant PDG Governance Groups and any other relevant group that will have oversight of the follow-up actions.

For activities that are jointly funded with another donor, if the co-funder funds the evaluation, MFAT must still publish the report and an accompanying management response.

Learning

Each evaluation report will have a clear and succinct presentation of key findings, lessons, and recommendations. Evaluators, in consultation with MERL Unit, will develop a plan for communicating and disseminating the findings and lessons to both internal and external audiences. The MERL Unit will produce regular syntheses of findings, lessons and results from evaluations, AMAs and ACAs, for PDG Plans and thematic areas. Evidence and lessons from evaluations will be used to inform business cases for new and extending activities, as well as inform PDG Plan Reflection Reports.

Policy review

The MERL Unit will review this evaluation operational policy annually. The MERL Unit will report to PDLT S&P with recommendations for improvements to this policy on an as required basis.

Success of this policy will be measured based on the extent to which evaluations are produced, used, and valued in informing strategy development, programme planning, learning and decision-making.

⁵ Completion of evaluation report is defined as MFAT acceptance of the final report.

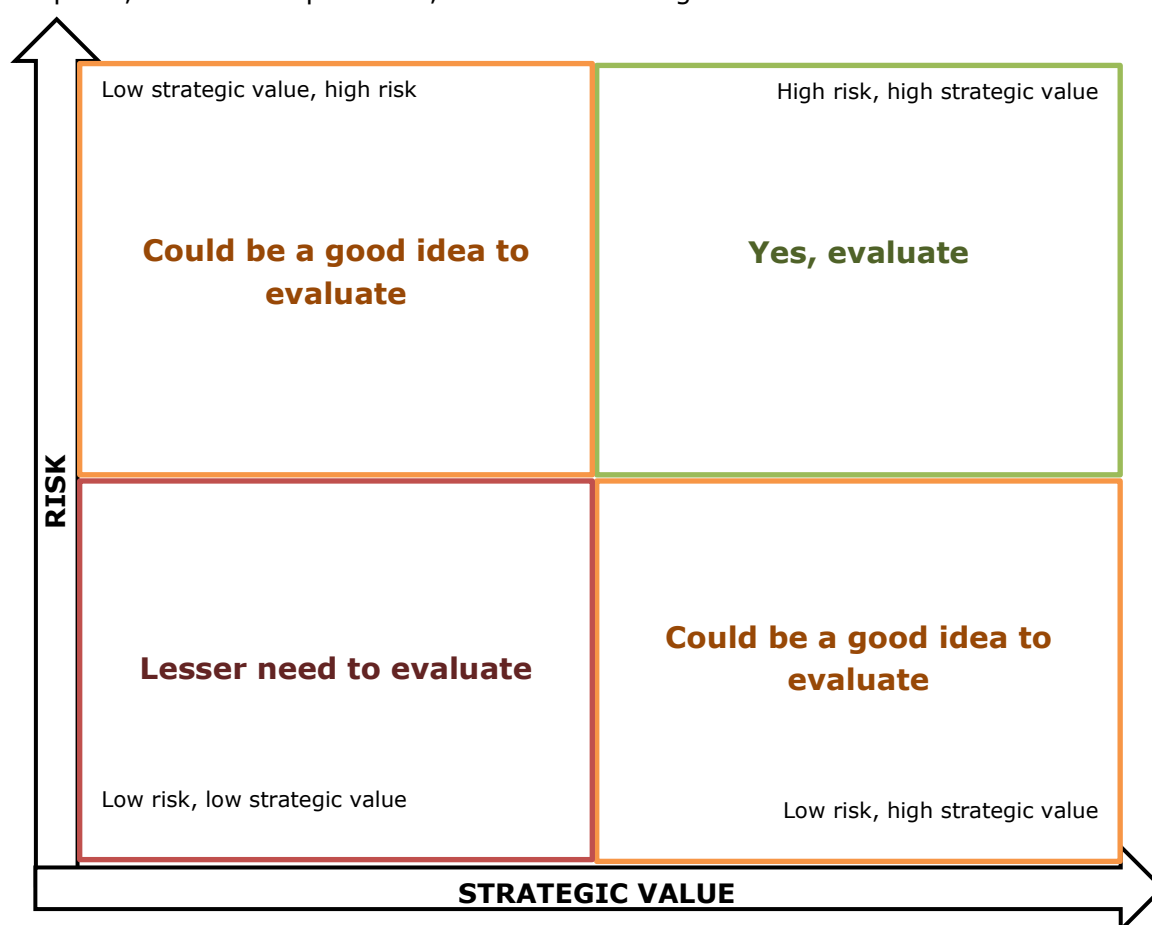
Appendix A: Evaluation prioritisation

The following guidance is provided to support governance groups when making decisions about whether to evaluate an activity during implementation, or after completion.

These considerations should be made during assessment of activity business cases and review of findings and actions from partner reports and activity monitoring and completion assessments (AMAs and ACAs). The diagram proposes a method to prioritise using **Risk** and **Strategic value** as the two main criteria.

Risk includes thinking about complexity of the activity (design features, context/environment, modality), the duration, and the size of investment.

Strategic value includes thinking about geo-strategy, bilateral importance, regional impacts, domestic imperatives, international obligations etc.



If you assess that your activity fits in the top right box, an evaluation should be undertaken. If your activity fits in the bottom left box then an evaluation is likely to not be necessary if there is sufficient monitoring in place supported by activity monitoring and completion assessments (AMAs and ACAs).

For example: a highly complex, dynamic activity expected to involve many uncertainties and issues, which is of high strategic importance, and has costs of \$20 million over four years implementation – this should be evaluated.