



Performance Assessment and Evaluation

Overview

The Australian aid program is committed to strengthening its performance orientation to help managers improve development effectiveness, and account for results. This policy sets out the minimum expectations for measuring performance at strategy or program level and at activity level. It describes three types of reporting: annual performance reporting, the quality reporting system and evaluation and review

Policy

Contents:

Introduction

1. Purpose
2. Policy overview

Principles

1. Clear objectives
2. Transparency
3. Contestability and sound evidence
4. Whole of government and other partnerships
5. Alignment and harmonisation
6. Efficiency
7. Appropriate mix of independence and self-assessment

The three types of performance reports

1. Annual Performance Reports
2. Quality Reporting
3. Evaluation

Annex: Performance terminology

Introduction

1. Purpose

The purpose of this policy is to continually improve performance and learning by assessing whether objectives are being achieved and remain relevant. Implementation of this policy also helps meet our accountability requirements to the Australian Parliament and public.

There are three reasons for assessing performance of the aid program: management, learning and accountability:

- Management

Implementing this policy will help program managers and partners to focus on results and improve quality by collecting reliable performance information. It will help managers to deliver against targeted results, promptly address what is not working well and inform program and budget decisions.

- Learning

Regularly reviewing the effectiveness of the aid program will provide a knowledge base for staff and partners to learn more about what is working well.

- Accountability

Reports produced under this policy will supply the Minister, Parliament, public and partners with information on the results achieved with Australian funds. Specifically, corporate accountability for quality is met by through the Quality Reporting System. Reports will be made publicly available as a rule,

helping to build broad support for the aid program. Reports will inform the Annual Review of Development Effectiveness produced by the Office of Development Effectiveness and can be used in preparing answers for Senate Estimates questions, in Government audit processes and to inform the DAC peer review and reporting against the Paris Declaration on aid effectiveness.

2. Policy overview

This policy sets out expectations for measuring the performance of Australian aid, including a number of principles that apply to all types of reporting. There are three types of reporting processes: annual performance reporting, the quality reporting system and evaluations and reviews. These are summarised in the table below.

Description of Performance Assessment Activities

Report	Coverage	Role	Primary purpose and audience
Annual Performance Report	Country and regional programs (Annual Program Performance Report)	Describes the performance of the program over the year, rating progress against objectives, using information and approaches defined in the Country or Regional Strategy Performance Framework.	Primarily for managers of the program, the Program Committee and partners to improve effectiveness of strategy or program implementation. Also serves accountability purpose and provides information for the ARDE.
	Sector and cross-sector programs (Annual Thematic Performance Report)	Reports against the objectives set in the relevant sector strategy or policy document, or the program design. Also provides reporting against new policy proposals framed by sector.	
	Global programs	Reports against individual AusAID developed performance frameworks that outline explicit objectives which underpin Australian support to the relevant program.	
Quality Reporting System	All monitored activities	Reports and rates each aid activity according to a set of common quality principles at three different stages – entry, implementation and completion.	Ensure new activities are of the highest quality before implementation, to enable continuous quality improvement during implementation and accounts for activities at completion. The tool is also used to generate annual statutory corporate performance information. Main audiences are program managers and Program Committee. OPMU will use the scores to draw out common areas of strength and weakness for corporate learning. ODE will use analysis of scores to assess quality of the aid program
Evaluation	Country strategies, monitored activities, sectors and themes	In depth assessments (often independent) of programs or activities, mainly focused on achievement of their objectives and their relevance.	Primarily for generating evidence for accountability and for management learning for future strategies or activities. Important information for the ARDE.

Principles

Some common principles emerge from Australian and international experience with evaluation and performance assessment. These principles apply to all forms of performance reporting outlined in the table above, and are:

1. Clear objectives

All aid interventions, whether an individual activity or country, regional or thematic program must identify objectives. These are the priority outcomes to be achieved by the end of the intervention. Objectives spell out the primary basis for judging success of the intervention. See annex for definitions.

2. Transparency

Performance of the aid program should be open and transparent to partners, beneficiaries and the public, both in Australia and partner countries. This can be achieved in a number of ways. Making all reports publicly available (while protecting the confidentiality of individual informants) is the default position.

3. Contestability and sound evidence

The rigour of reporting on performance is enhanced when the findings stand up to scrutiny and challenge, by management, peers and external individuals. As a general rule, all aspects of performance reporting should be subject to contestability.

There should be a sound basis of evidence supporting the conclusions drawn in performance reporting and evaluation. This presents a challenge in countries where data is scarce, out of date or unreliable. Full use of both quantitative and qualitative information needs to be made in forming judgements. Country, regional and thematic programs should plan for regular evaluations, including independent completion reports.

4. Whole of government and other partnerships

Ideally performance of the aid program would be routinely considered in ongoing conversations with partner governments, other donors or major partners such as other Australian government departments. The degree to which partners should be involved in assessing performance should be considered case by case: with major partners, consideration should be given to the value of undertaking joint assessments. When preparing reports, staff should seek input and consult with their most significant government department partners. As a minimum, performance reports should be shared with all major partners.

5. Alignment and harmonisation

Aligning performance requirements with those of partner countries and relying as far as possible on partner countries' own reporting frameworks is good practice. Country and regional programs should work with other donors to identify and develop partner government information and review systems which will inform progress towards common objectives. Where necessary, programs may use their own systems for gathering and analysing information – using a harmonised and cost effective approach that could be adopted by partner governments at a future date. The same principles apply at activity level, through activity monitoring and evaluation systems.

6. Efficiency

The amount of effort and resources invested in the reporting process, including collecting and analysing performance data, should be proportional to the value and nature of the program.

7. Appropriate mix of independence and self-assessment

Independence is important in terms of credibility to an external audience, while self-assessment can be most valuable in terms of management utility of the findings. It is important to select the appropriate degree of independence to suit the purpose of a particular performance assessment or evaluation and be explicit about this in the report.

The three types of performance reports

There are three types of performance reports: Annual Performance Reports, the Quality Reporting System and Evaluations (see Box 1). These are described in the following sections.

1. Annual Performance Reports

This section describes reporting at a strategy or program level. Reporting is against a standard set of three questions:

- what are the *results* of our aid program?
- what is the *quality* of our aid activities?
- what are the *management consequences* of this assessment?

1.1 Country and regional programs

Country (and regional) programs are the cornerstone of the aid program. Whole of Government strategies describe specific development objectives to be achieved and include operational performance frameworks which capture information on an ongoing basis to help the production of *Annual Program Performance Reports*, sector and thematic reporting and evaluation of the strategies. In total, country and regional programs account for around two thirds of total Australian aid spending.

The amount of effort that is put into reporting should be proportional to the value and nature of the program. The level of seniority in peer reviews of draft reports would also differ proportionately.

1.2 Global Programs

For the purposes of reporting, non-country programs include the following areas. They account for the remaining third of Australian aid spending:

- Core funding to multilateral organisations, including the World Bank; ADB and other international financial institutions, as well as UN and other multilateral agencies
- Scholarships
- Community and Business Partners
- Humanitarian and Emergencies

1.3 Thematic areas

While the bulk of aid expenditure is covered through the above reports, looking at performance through a sectoral or cross sectoral lens is important for understanding the progress we are making. This includes areas of governance, infrastructure, environment, health and education.

Besides being a way to capture progress against policies (where they exist), the *Annual Thematic Performance Report* is also the mechanism for reporting against new policy proposals. Each year, any thematic area that has spent new funding prepares an *Annual Thematic Performance Report*, covering both existing and new money. Broad areas (for example governance) may choose to focus their annual report on a particular theme.

2. Quality Reporting

The system for reporting on the *quality* of aid activities supports program management and adds to accountability of funds committed for specific aid objectives. It also provides the data to answer the second question in Annual Performance Reports on the quality of Australia's aid interventions. The tool is used to generate statutory annual corporate performance information.

2.1 Aid Quality Principles

Quality is the extent to which aid activities apply internationally recognised characteristics of good aid practice, summarised in the five quality principles.

All activities are expected to:

- Achieve clearly stated objectives that contribute to higher level objectives in the program strategy
- Effectively measure progress towards meeting objectives

- Continually manage risks
- Appropriately address sustainability, with due account of partner government systems, stakeholder ownership and phase out
- Be based on sound technical analysis and continuous learning

2.2 Monitored activities

Quality reporting at entry, during implementation and at completion is mandatory for all substantive, monitored aid activities, across all aid modalities, eg, cofinancing, sector-wide approaches, global programs and bilateral activities, where:

- the expected Australian Government funding over the entire life of the activity is greater than \$3 million or
- the value is less than \$3 million, but the activity is significant to country or corporate strategies or key relationships with other development partners including whole of government partners.

2.3 Quality At Entry

The design of all new monitored activities goes through a quality enhancement process comprising three steps – a concept peer review, an independent appraisal and an appraisal peer review. A Quality at Entry Report is the product of the appraisal peer review, chaired by either the relevant Minister Counsellor or ADG and attended by an appropriate mix of internal and external expertise to maximize opportunities for improvement and ensure the contestability of the design. If a peer review finds that the design of an activity does not meet AusAID's quality principles, the Minister Counsellor or ADG may require amendments to the design and a further Quality at Entry report.

2.4 Quality at Implementation

Quality at Implementation reports will be completed at least once a year, typically in preparation for the Annual Performance Report. Updates may also be prepared ahead of independent reviews or management events such as Technical Advisory Groups, Project Coordinating Committees and Mid-Term Reviews, and revised as needed to reflect the outcomes of such reviews. All ratings will be entered by the activity manager and approved by the line manager.

The Quality at Implementation ratings will be the responsibility of the program team, but can be contested by comments from independent reviews, OPMU, Thematic Groups, ODE and others. This could lead to a new assessment. More frequent reviews may be called for on activities with marginal or unsatisfactory ratings, including those facing significant risks. Relevant management staff (e.g. Assistant Director General or Minister Counsellor) are responsible for ensuring that ratings are reviewed and moderated across the portfolio of activities in each program.

2.5 Quality at Completion

All activities must have some form of completion report. For monitored activities this is fulfilled by an Independent Completion Report which includes quality ratings (see evaluation).

3. Evaluation

Evaluations generate evidence to demonstrate effectiveness and provide management with lessons to make informed decisions about the future. Analysis and data from evaluations help answer the question in the Annual Performance Report on the results of Australia's aid interventions.

The terms evaluation and review are sometimes used synonymously. The main point of difference is that 'evaluation' describes a more comprehensive or in depth assessment. This policy covers both evaluations and reviews, such that references to evaluation also cover review.

3.1 Definition and scope of evaluation

Evaluations are *assessments, as systematic and objective as possible, of ongoing or completed aid activities, programs or policies, their design, implementation and results.*

The main focus of evaluations (at this stage) is on determining the extent to which objectives have been achieved and their relevance. They may also look at efficiency, impact and sustainability (see Annex).

Evaluation should be integrated into country, regional and sectoral strategies. Each Country or Regional Strategy Performance Framework, a key component of a strategy, should include planned evaluations. In addition to the principles outlined in section 3, the DAC Evaluation Quality Standards should be consulted and, to the extent possible, applied in all evaluations.¹ Among other things, these include adopting a gender sensitive approach and ensuring that the evaluation process helps to build evaluation capacity in partner countries.

Harmonisation and alignment principles apply to the conduct of evaluations. Collaboration with partner country representatives, other donors and research institutions in joint evaluations is encouraged, and should increase over time.

3.2 Types of Evaluation

The main types of evaluations necessary for the aid program include activity level evaluations, country or regional strategy reviews and thematic or sector evaluations. The following sections give some information about three types of evaluation:

3.2.1. Activity Level Evaluations

These evaluations focus on assessing the performance of a specific activity. The primary evaluation at the activity level is the Independent Completion Report. These are mandatory for all monitored activities. ICRs should be independent. Other types of evaluations conducted at the activity level include Mid-Term Reviews and ex-post impact evaluations. Depending on the primary purpose of these evaluations, program teams should exercise their judgment to determine the right balance between independence and program participation, and ensure the degree of independence of the evaluation or review is documented in the report.

3.2.2. Country/Regional Strategy Reviews

Strategy reviews focus on the entirety of Australian ODA to a partner country or region. Strategy reviews are mandatory. They must be undertaken in the lead up to the development of new strategies. They should be conducted by people independent of the program over the life of the strategy.

3.2.3. Cluster, thematic or sectoral evaluations

Thematic evaluations focus on selected aspects or themes common to numerous development activities (e.g. cross-cutting or overarching themes, such as governance, gender and partnership approaches for aid delivery or policies regulating the aid program). Cluster evaluations assess a group of similar activities. Depending on the primary purpose of these evaluations, program teams should exercise their judgment to determine the right balance between independence and program participation, and ensure the degree of independence is documented in the report.

¹ The DAC Evaluation Quality Standards have been approved by the members, including Australia, of the DAC Network on Development Evaluation. The standards are available at: www.oecd.org/dac/evaluationnetwork.

Annex: Performance terminology

Introduction

This policy supports application of international criteria for evaluating development assistance, namely relevance, effectiveness, efficiency, impact and sustainability². However, during this first phase of policy implementation it is impractical to expect all the criteria to be addressed with equal weight, and the emphasis has been placed on measuring relevance and effectiveness i.e. establishing clear objectives and measuring achievement of them.

The purpose of this annex is to provide clarity around a minimum set of performance terms that will help us do that (the DAC glossary can help in providing definitions of the full range of performance terms³). Confusion about the meaning of particular words is by no means the most significant challenge we face in implementing the performance agenda, but it can affect how we think about performance, how effectively we use new corporate tools for monitoring and reporting, and the value we get from training and guidance in this field.

The terms addressed here are: Objective; Outcome; Output; Activity

Definitions

Objective:	A specific statement setting out what an activity or other intervention is expected to achieve, by a given time.
<p>All AusAID activities must identify their objectives. We define an objective to be:</p> <ul style="list-style-type: none"> (i) an explicit statement (i.e. written down); (ii) an outcome; (iii) what we aim to achieve by the end of the activity; and (iv) measurable – managers will know and can demonstrate the extent to which it has been achieved. <p>The extent to which objectives are achieved, therefore, provides the primary basis for judging the success of any activity at its end point. Managers are not accountable for the objective, but they are accountable for managing towards it.</p> <p>Good management practice means that we should limit the number of objectives so that we focus on the priority outcome(s) to be achieved.</p> <p>A challenge for us is pitching objectives at the right level – i.e. selecting the appropriate outcomes to aim for. The decision will be influenced by the relative scale and timeframe of our support in addition to how well placed we are to influence change. Circumstances will vary and so will the level at which we pitch our objectives. <i>The guiding principle in this is to select the highest possible outcome that we can plausibly influence within the timeframe of our support.</i></p> <p>Interventions should also identify what their objectives contribute to in terms of higher-order outcomes. These are beyond the scale of the objectives we have set for our assistance, but our objectives actively contribute to them. In practice, the higher-order outcome for an activity would usually be the relevant country strategy objective, but it could be an objective under one of our thematic policies for the aid program. For a country strategy objective, it might be the relevant national development outcome targeted in the partner country's development plan. The main point being that we need to be clear about how an intervention's objective(s) contribute to the broader development agenda.</p> <p>In practice, we cannot wait for the end of the activity to know if we are being successful. Our program logic will therefore usually include more immediate outcomes (see below) which we will track during implementation. These may be components which taken together will contribute to the objective, or a set of changes along the way to the objective. In fluid and complex programs, where progress will not be linear, managers will need to revisit their performance expectations annually.</p>	

² See the *DAC Principles for the Evaluation of Development Assistance*, OECD (1991)

³ *Glossary of Terms Used in Evaluation, in 'Methods and Procedures in Aid Evaluation'*, OECD (1986), and the *Glossary of Evaluation and Results Based Management (RBM) Terms*, OECD (2000).

Outputs:	Products, goods or services delivered by an activity or other intervention
<p>Outputs are what our activities produce or our money pays for. Examples include facilities built, training programs run, workshops held, analyses/reports produced. They are within the control of a particular intervention – if not entirely then largely, although they may be jointly delivered with partners. Although they can be significant, outputs are not ends in themselves; they are the means to influence (positively) the outcomes we are seeking to achieve.</p>	

Outcomes:	Intended or unintended effects of the outputs from an activity or other intervention.
<p>Outcomes are the events or changes in conditions, institutional arrangements, behaviour or attitudes that we hope will occur as a result of our intervention. (We should also recognize the possibility of unintended or negative outcomes occurring). In contrast to outputs, outcomes are the results from the actions of multiple stakeholders but they are influenced by the outputs of the intervention. Example outcomes could include: improved exam results for school children; reduced delays and fraudulent losses in public procurement; the adoption of improved management arrangements within primary health care systems.</p> <p>In practice, the difficulty is not identifying outcomes but knowing where to pitch our objective among the possible outcomes in a way that will realistically drive management for results (if too low we are too easily satisfied with insignificant results; if too high managers can ignore the objective or give up trying to meet it).</p> <p>When thinking about outcomes, it can be helpful to distinguish between ‘immediate’ or short-term outcomes (e.g. levels of satisfaction with the training among people who attended a smoking cessation course), ‘intermediate’ or medium-term outcomes (e.g. numbers of people still not smoking 12 months after completing the course), and ‘ultimate’ or long-term outcomes (e.g. reduction in death rate from smoking-related diseases). Progressing from immediate towards longer-term outcomes takes us closer to the real reason for providing support, but we also get further away from our influence.</p>	

Main Implications

We should use the term ‘activity’ to describe any unit of AusAID work with a budget and authorised spending. This will help resolve ongoing confusion about how we refer to projects, programs, facilities and other interventions. In practice, existing systems (notably AidWorks) will not be entirely consistent with the new definitions proposed. We will be working to improve consistency over time.

We need to be strict when we use the term ‘objective’. AusGuide used to encourage use of the term at any level of the log-frame hierarchy from output upwards. This is no longer the case. To embed this change within AusAID and with our partners, we should where necessary stop using the term “component objectives” and use “intermediate outcomes” instead. We should also stop using the term “purpose” (for example in a logframe) and use “objective” as directly analogous.

Designers and managers need to clearly distinguish between outputs and outcomes.