



Australian Government  
Productivity Commission

# A Guide to Evaluation Under the Indigenous Evaluation Strategy **Draft**

The Commission acknowledges the Traditional Owners of Country throughout Australia and their continuing connection to land, waters and community. We pay our respects to their Cultures, Country and Elders past and present.

© Commonwealth of Australia 2020



Except for the Commonwealth Coat of Arms and content supplied by third parties, this copyright work is licensed under a Creative Commons Attribution 3.0 Australia licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/3.0/au>. In essence, you are free to copy, communicate and adapt the work, as long as you attribute the work to the Productivity Commission (but not in any way that suggests the Commission endorses you or your use) and abide by the other licence terms.

### Use of the Commonwealth Coat of Arms

Terms of use for the Coat of Arms are available from the Department of the Prime Minister and Cabinet's website: <https://www.pmc.gov.au/government/commonwealth-coat-arms>.

### Third party copyright

Wherever a third party holds copyright in this material, the copyright remains with that party. Their permission may be required to use the material, please contact them directly.

### Attribution

This work should be attributed as follows, *Source: Productivity Commission, A Guide to Evaluation under the Indigenous Evaluation Strategy, Draft*.

### An appropriate reference for this publication is:

Productivity Commission 2020, *A Guide to Evaluation under the Indigenous Evaluation Strategy, Draft*, Canberra, May.

### Publications enquiries

Media, Publications and Web, phone: (03) 9653 2244 or email: [mpw@pc.gov.au](mailto:mpw@pc.gov.au)

### The Productivity Commission

The Productivity Commission is the Australian Government's independent research and advisory body on a range of economic, social and environmental issues affecting the welfare of Australians. Its role, expressed most simply, is to help governments make better policies, in the long term interest of the Australian community.

The Commission's independence is underpinned by an Act of Parliament. Its processes and outputs are open to public scrutiny and are driven by concern for the wellbeing of the community as a whole.

Further information on the Productivity Commission can be obtained from the Commission's website ([www.pc.gov.au](http://www.pc.gov.au)).

---

# Contents

|  |    |
|--|----|
| A Guide to Evaluation under the Indigenous Evaluation Strategy | 1  |
| 1 Building evaluation into policy and program design           | 4  |
| 2 Deciding what policies and programs to evaluate              | 10 |
| 3 Evaluation planning, design and conduct                      | 14 |
| 4 Reporting evaluation findings                                | 31 |
| 5 Building capability and a culture of evaluation              | 34 |



---

# A Guide to Evaluation under the Indigenous Evaluation Strategy

The Indigenous Evaluation Strategy (the Strategy) provides a whole-of-government framework for Australian Government agencies to use when selecting, planning, conducting and using evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. The Strategy puts Aboriginal and Torres Strait Islander people at its centre, and recognises that governments need to draw on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people if outcomes are to improve.

## **A new approach to evaluation: an Indigenous Evaluation Strategy**

For decades there have been calls to better understand how policies and programs are affecting Aboriginal and Torres Strait Islander people. However, there continues to be limited evidence about the effectiveness of many policies and programs designed to improve outcomes for Aboriginal and Torres Strait Islander people. Wide gaps between Aboriginal and Torres Strait Islander people and non-Indigenous people also persist across many indicators of wellbeing — life expectancy, health, educational attainment and employment — notwithstanding more than ten years of Closing the Gap initiatives.

Evaluation — the systematic assessment of a policy or program’s design, implementation and outcomes — is about understanding what governments are doing right, what they are getting wrong and where they can do better. The Strategy is an opportunity to improve the quality and use of evaluations of Australian Government policies and programs affecting Aboriginal and Torres Strait Islander people, and to provide insights more generally on how to assess complex policies and programs affecting all Australians.

To achieve better policy outcomes, Aboriginal and Torres Strait Islander people need to be at the centre. What they value, their expertise and lived experience should be reflected in what is evaluated, how evaluation is undertaken and the outcomes policies seek to achieve. If the outcomes of policies are not what is valued by Aboriginal and Torres Strait Islander people, then the policies have limited value and little prospect of success.

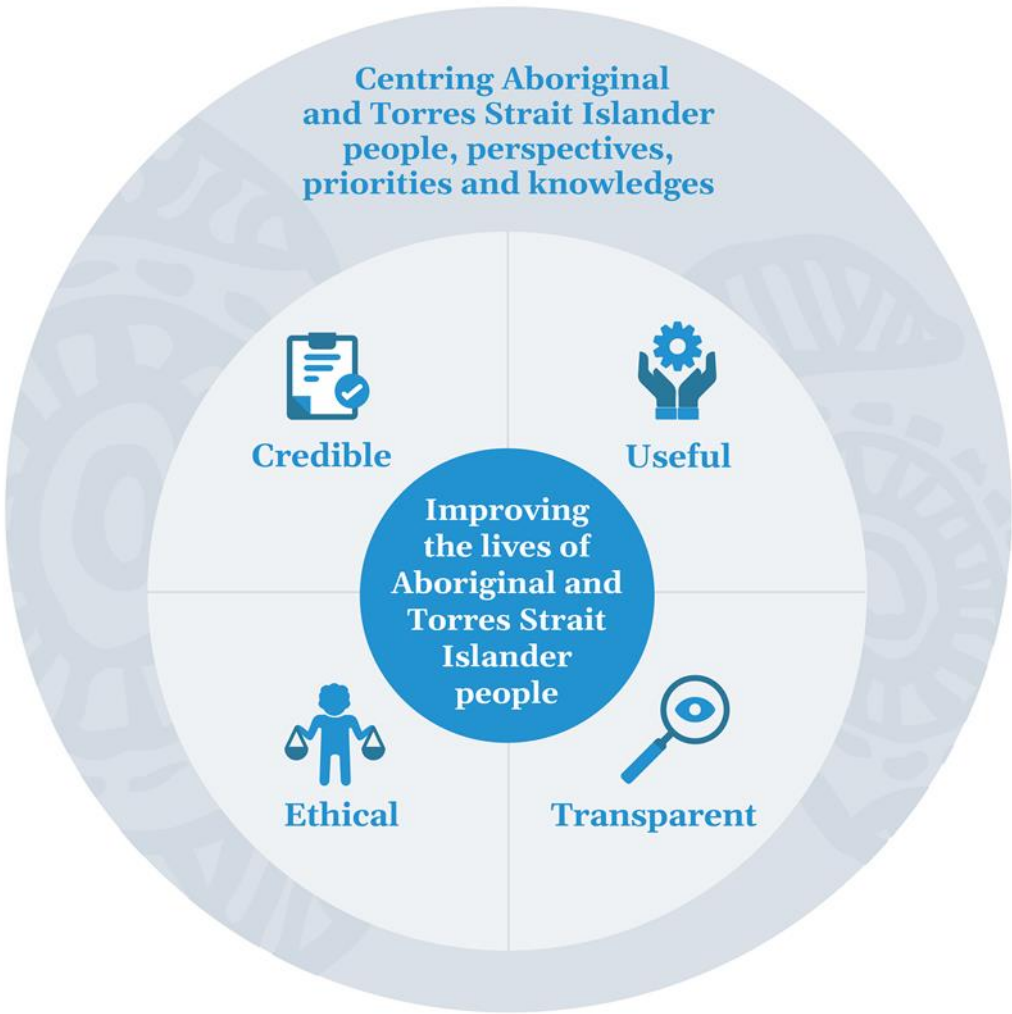
As well as providing guidance to agencies on planning, conducting and using evaluation, the Strategy needs to support more effective ways of working with Aboriginal and Torres Strait Islander people, and include governance arrangements that embed incentives to learn from and use evaluation findings in policy decision making. It is better policies, not evaluation per se, that will improve the lives of Aboriginal and Torres Strait Islander people. This is a

fundamental shift from the way evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are currently undertaken.

### Guiding principles for the Strategy

The overarching principle of the Strategy is centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (figure 1). This principle is about recognising the strengths of Aboriginal and Torres Strait Islander people, communities, knowledges and cultures. It is also about building genuine partnerships with Aboriginal and Torres Strait Islander people to define policy and program outcomes, and decide on evaluation questions, how evaluations will be conducted and how evaluation findings will be interpreted.

Figure 1      **Guiding principles for the Indigenous Evaluation Strategy**



---

The overarching principle is also the lens through which the Strategy’s other principles — credible, useful, ethical and transparent — are interpreted. These principles frame how agencies should plan and conduct evaluations and how evaluations will be assessed.

## **This guide provides practical advice on implementing the Strategy**

This guide has been developed as a companion to the Strategy. It provides practical advice for Australian Government agencies on how to conduct evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

The guide does not require users to have in-depth evaluation and methodological expertise. Rather, it is written for program managers and those commissioning and using evaluations. Links to further evaluation references, and more specialist sources, are provided throughout the guide.

The guide steps through the different stages of evaluating policies and programs and building a culture to support evaluation under the Strategy:

- building evaluation into policy and program design
- deciding what policies and programs to evaluate
- evaluation planning, design and conduct
- reporting evaluation findings
- building capability and a culture of evaluation.

At each of these stages, the guide outlines the questions to consider to ensure that commissioners and users of evaluation respond to the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people and integrate them into evaluation processes to improve the quality and usefulness of evaluations.

The Strategy does not replace, but rather complements and builds on Australian Government agencies’ existing evaluation processes and frameworks. It provides an overarching whole-of-government framework and consistent accountabilities for all Australian Government agencies developing and implementing policies and programs that affect Aboriginal and Torres Strait Islander people.

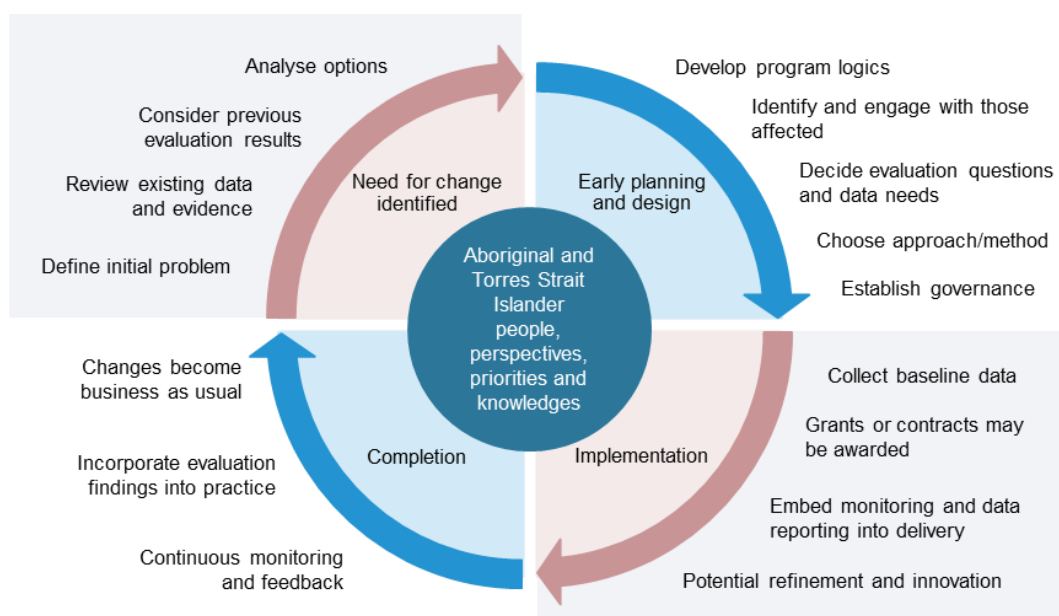
---

## 1 Building evaluation into policy and program design

Evaluation is most effective when it is planned for early and when it is integrated into each stage of policy and program design — from setting policy objectives through to using evaluation findings to inform future policies and programs. Further, in the context of policies and programs covered by the Strategy, it is critical that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges be carefully considered at each of these stages (figure 2).

---

Figure 2 **Evaluation should be considered at every stage of policy and program design and implementation**



Before a policy or program is implemented, evaluation should inform policy design. While a policy or program is being implemented, evaluation should inform thinking on ways to improve the policy or program. And after the policy or program has been implemented, evaluation can provide evidence on outcomes and impacts that can be drawn on to inform future policies and programs.

### Drawing on evaluation evidence when designing a policy or program

Good practice policy and program design draws on existing evidence about what works and why, including evidence from previous evaluations. Policy makers should examine the existing evidence base at the earliest stages of policy and program design to see what lessons there are from similar policies or programs. This includes assessing the strengths (and weaknesses) of the evidence supporting the assumptions upon which a policy or program is based. It also involves looking at the evidence from trials or pilots of similar policies or



---

programs and relevant behavioural insights on the best way to implement the policy or program and what the likely outcomes could be.

When assessing existing evidence, it is important to critically examine the strengths and limitations of the evidence and check that it is relevant for the proposed policy or program situation, setting and local context. An assessment of the existing evidence can also highlight gaps or unanswered questions that evaluation could seek to address (box 1).

### **Box 1            Review and synthesis of existing evidence**

Working out what is already known about a particular form of intervention is not always easy, particularly when there are volumes of research evidence. And because not all evidence is of equal value, some way of differentiating between high and low-quality studies is needed.

Evidence synthesis methodologies aggregate evaluation findings and review them in a systematic way (and generally also assess and rate the strength of the evidence).

Systematic reviews aggregate results that fit a pre-specified eligibility criteria in order to answer a specific research question. They aim to minimise bias by using explicit, systematic methods that are documented in advance with a protocol.

Systematic reviews search all the available research evidence for a particular question, which counters problems of selection bias that come from only identifying studies that are readily accessible or only published on major databases. They can help highlight gaps in the evidence (and identify evaluation questions that need to be answered) and methodological issues that may need to be considered in the design of an evaluation.

A meta-analysis pools statistical results from multiple studies as a way to increase explanatory power and improve estimates of the size of effects. A number of organisations collate and synthesise research and evaluation evidence relating to Aboriginal and Torres Strait Islander people. The Australian Indigenous HealthInfoNet, for example, provides an evidence base to inform practice and policy in Aboriginal and Torres Strait Islander health. The Lowitja Institute also publishes all its research on the Institute's website.

The Australian Institute of Health and Welfare is developing an Indigenous Community Insights website (a one-stop-shop for data and statistics about Aboriginal and Torres Strait Islander people's health and wellbeing) and an Indigenous Mental Health and Suicide Prevention Clearinghouse. An Indigenous Research Exchange is also being developed by the Australian Institute of Aboriginal and Torres Strait Islander Studies.

---

### Questions to consider when drawing on evidence when designing a policy or program

- Is there evidence from other policies and programs on the problem or issue you are seeking to address? Does it validate your hypothesis and drivers of change?
- How credible and relevant is the evidence to support or validate the success factors or outcomes you are seeking? Have differences in context and the strengths and limitations of the existing evidence been considered? Does the evidence shed light on any potential unintended effects or behavioural changes from similar policies or programs?
- Has the evidence been tested in Australia or in a similar setting and/or a similar population? Is there evidence demonstrating the effects for Aboriginal and Torres Strait Islander people?
- Has evidence from a range of sources been examined, including insights from Aboriginal and Torres Strait Islander people and organisations?
- Does the evidence provide any guidance (resources, timing, capability) on how to implement the policy or program?

### Planning for evaluation during policy and program design

Planning early and being clear about the objectives of a policy or program can result in both better policies and better evaluation.

Developing an evaluation plan as part of the policy or program design has a number of benefits. It can help ensure:

- evaluations identify the objectives and outcomes of a policy or program and clarify what the evaluation questions should be
- the right data are collected (baseline and monitoring data) for the evaluations and data collection systems are in place from the start — this will improve the quality, rigour and usefulness of the evaluation
- agencies consider whether policies or programs can be designed or implemented in ways that support data collection and particular evaluation approaches
- Aboriginal and Torres Strait Islander people, and others affected by a policy or program, are engaged with early on in the process and have an input into policy objectives and outcomes as well as evaluation questions, approaches, methods and how evaluation results will be reported and disseminated
- the different effects of mainstream policies and programs on Aboriginal and Torres Strait Islander people, and potentially other minority groups, are considered (and the data required to evaluate the effects are collected)
- adequate resources and realistic timeframes are allocated for the evaluation.

Developing a *program logic* (also called a logic model or theory of change) is a critical first step in evaluation planning. A program logic describes the relationship between policy or

---

program inputs, outputs and outcomes, and spells out the assumptions that underpin the relationships. It is a useful way to be explicit about the evidence you are using, how you expect to get results, what outcomes you will measure, and how you will measure the outcomes.

Under the Strategy, all New Policy Proposals that are expected to have significant and/or differential impacts on Aboriginal and Torres Strait Islander people are to include an Indigenous Evaluation Threshold Assessment (box 2). This will ensure that these policies and programs have a preliminary evaluation plan and budget in place.

### **Box 2                    An Indigenous Evaluation Threshold Assessment**

For those new policies and programs that will have a significant and/or differential effect on Aboriginal and Torres Strait Islander people, the New Policy Proposal should include an Indigenous Evaluation Threshold Assessment. The assessment should include details on:

- (a) the Aboriginal and Torres Strait Islander people, organisations and communities that would be affected by the proposed initiative
- (b) a proposed engagement plan (including timeframes and cultural safety considerations)
- (c) a proposed approach (including the scale of evaluation required)
- (d) the data required to assess the policy's impact and how it would be collected
- (e) an estimated evaluation budget.

### **Evaluation in practice: Budgeting for evaluation during policy and program development**

A new Indigenous ranger program is being developed for roll out in remote areas. Because the program will be targeted at Aboriginal and Torres Strait Islander people, and could have a significant effect on their wellbeing, an Indigenous Evaluation Threshold Assessment is being prepared as part of the New Policy Proposal. The plan is for the evaluation to include quantitative analysis of administrative data as well as qualitative case studies in several locations where the program is operating.

When estimating the cost of the future evaluation, the agency has budgeted extra time and money (compared with an earlier evaluation of a similar program rolled out in regional areas) on the basis that travel costs are likely to be higher, and evaluators will need to visit each case study area several times to build up a relationship with community members. In some areas where the program is proposed to run, English is spoken as a second or third language by many people, and the evaluation budget has allowed for translation of surveys and responses.

The proposed timeline for the evaluation has also factored in the need for a formal ethics review, as well as meeting the requirements of research protocols that a number of communities likely to be involved in the case studies have in place.

---

## Policy design and evaluation

At the early stages of policy and program design, agencies should think about whether the policy or program can be designed or implemented in a way that aids evaluation. Rigorous impact evaluation requires an estimate of the counterfactual: what would have occurred in the absence of the policy or program? If designed well, piloting, randomisation and phased introduction of a policy or program can help evaluators to compare the outcomes of a policy or program on participants with those of non-participants.

The use of pilots and trials can also allow policy makers to experiment with policies and programs on a small scale. By building evaluation into a pilot, lessons on implementation and impact can be used to refine a policy or program before it is rolled out more widely.

### **Evaluation in practice: Lessons from a pilot program**

A new legal assistance program targeted at young people is being developed. It uses an innovative approach to provide integrated legal and other support services for young people at risk of incarceration. Because this approach has not been tried before in Australia, the agency wants to introduce it as a pilot program before rolling it out across Australia. The aim is to see if the approach improves outcomes, as well as iron out any implementation issues if it is to be implemented more widely.

To aid rigorous evaluation, the sites for the pilot are chosen to ensure that a well-matched control group of non-participating sites can be used for comparison. Both a process evaluation — focusing on whether the program is being implemented as planned — and an impact evaluation — to be conducted at the conclusion of the pilot — are planned. The pilot sites are chosen to reflect the diversity of settings in which the program might eventually operate, as well as to allow inferences to be made about the impacts on key participant groups, including Aboriginal and Torres Strait Islander people.

The agency uses the results to decide whether the program should be expanded, and if so, how it should be implemented to maximise positive outcomes. As a key objective of the pilot is to learn whether this innovative approach is suitable for use in Australia, the agency ensures that the lessons from the evaluation are shared widely.

---

## Questions to consider when planning for evaluation during policy and program design

- Will the policy or program have a significant and/or differential effect on Aboriginal and Torres Strait Islander people? If so, has an Indigenous Evaluation Threshold Assessment been undertaken?
- Are the program objectives clearly identified and has a program logic or theory of change been developed? What are the gaps in the existing evidence base that evaluation of the proposed policy or program could fill?
- Does the program logic consider the impact of the policy or program on Aboriginal and Torres Strait Islander people?
- Does the program logic focus on outcomes that are important for the lives of Aboriginal and Torres Strait Islander people?
- Could there be unintended negative consequences of the policy or program for Aboriginal and Torres Strait Islander people's wellbeing that are not captured in the program logic?
- Does the preliminary evaluation plan identify opportunities for meaningful engagement with Aboriginal and Torres Strait Islander people, and other stakeholders, during the evaluation?
- Does the proposed policy or program budget include sufficient resources for future evaluation, including resources to collect data and engage with stakeholders when developing an evaluation plan and undertaking evaluation? Are the time and resources allocated to evaluation proportionate to the size and importance of the policy or program being evaluated?
- Can the policy or program be designed or implemented in a way that aids evaluation (such as by piloting, randomly allocating participants and non-participants or a phased introduction)? Does the evaluation design allow for inferences to be made about how a pilot program might work in different contexts or for different groups of participants?

## Further reading on building evaluation into policy and program design

*Indigenous Evaluation Strategy, Draft Background Paper*, chapters 5 and 7.

HM Treasury 2020, *The Magenta Book: Central Government Guidance on Evaluation*.

---

## 2 Deciding what policies and programs to evaluate

It is not practical or feasible to rigorously evaluate every policy or program affecting Aboriginal and Torres Strait Islander people, so agencies need to decide what to evaluate. Setting evaluation priorities can help ensure evaluation resources are used in the most effective way to improve the lives of Aboriginal and Torres Strait Islander people.

When deciding what to evaluate, an agency should:

1. identify which of their new and existing policies and programs fall in scope of the Indigenous Evaluation Strategy, including those that contribute to government-wide evaluation priorities (table 1). For existing policies, agencies should undertake a stocktake of policies and programs to determine which fall within scope of the Strategy
2. use a criteria-based priority setting process (box 3) to determine which of the policies and programs within scope of the Strategy should be prioritised for evaluation (and the extent to which they should be evaluated).

---

**Table 1 Proposed interim government-wide evaluation priority areas**

Based on draft priorities established by the Council of Australian Governments and the Joint Council on Closing the Gap

### Priority Policy Areas

*Priority 1: Families, children and youth:* Aboriginal and Torres Strait Islander children thrive in their early years, and Aboriginal and Torres Strait Islander families and households are safe

*Priority 2: Health:* Aboriginal and Torres Strait Islander people enjoy long and healthy lives, and Aboriginal and Torres Strait Islander children are born healthy and strong

*Priority 3: Education:* Aboriginal and Torres Strait Islander students: succeed at school; stay in school; and reach their full potential through further education pathways

*Priority 4: Economic Development:* Aboriginal and Torres Strait Islander youth are engaged in employment or education, and strong Aboriginal and Torres Strait Islander workforce participation

*Priority 5: Housing:* Aboriginal and Torres Strait Islander people secure appropriate, affordable housing as a pathway to better lives

*Priority 6: Justice, including youth justice:* Aboriginal and Torres Strait Islander people are not overrepresented in the criminal justice system

*Priority 7: Land and waters:* Aboriginal and Torres Strait Islander peoples' land, water and cultural rights are realised

### Cross System Priority

Addressing racism, discrimination and social inclusion, healing and trauma, and the promotion of cultures and languages for Aboriginal and Torres Strait Islander people

### Priority Reform Areas

- Developing and strengthening structures to ensure full involvement of Aboriginal and Torres Strait Islander people in shared decision making
  - Growing Aboriginal and Torres Strait Islander community controlled services
  - Improving mainstream service delivery to Aboriginal and Torres Strait Islander people
  - Improving access to local data for Aboriginal and Torres Strait Islander people (proposed)
-

---

**Box 3****Using multi-criteria assessment to determine agency evaluation priorities**

Agencies should develop a fit-for-purpose **multi-criteria assessment process** to help categorise the relative significance of policies and programs in scope of the Strategy. To ensure consistency across agencies around prioritisation decisions, the criteria used for determining evaluation priorities should include:

- **Impact of the policy or program** — how likely the policy or program will impact the lives of Aboriginal and Torres Strait Islander people (taking into account the total Aboriginal and Torres Strait Islander population affected and/or whether a policy or program disproportionately affects Aboriginal and Torres Strait Islander people).
- **Strategic significance** — how important and valuable the policy or program is in terms of the competing priority areas within government (particularly government-wide evaluation priorities under the Strategy).
- **Risk profile** — the policy or programs overall risk. This may include how difficult it is to estimate the impact of the policy or program (the outcomes are uncertain or difficult to measure); or the risk that a policy or program could disproportionately affect the lives of Aboriginal and Torres Strait Islander people.
- **Program expenditure** — the overall funding for the policy or program. This could be measured in absolute terms, or relative to the cost of other programs delivered or managed by the agency.

Each of the criteria should be underpinned by the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people.

To determine a ranking of policies and programs for evaluation, agencies should:

- assign scores for each policy or program against each criterion above, based on defined thresholds
- calculate a weighted average score for each policy or program against the criteria to determine where it sits along a continuum of overall 'significance'. ('Weights' could be applied to all criteria to ensure outcomes are not biased against criteria that may be less important across policies and programs within an agency, or where the criteria may be less reliably measured.)

Based on resources and the level of activity, agencies would then establish a threshold so that all 'significant' policies and programs would be resourced to be rigorously evaluated. Lower priority policies or programs could focus their evaluative activity on improving data collection towards more meaningful outcome measures, monitoring and performance audits, or other evaluative activities that facilitate learning by doing, such as user surveys or reviews.

## Evaluation in practice: Balancing impact, risk, expenditure and strategic significance when determining evaluation priorities

An agency is applying a multi-criteria assessment to prioritise three programs for evaluation: a pilot intensive family support program being trialled in a remote community; a large mainstream education program; and an established Indigenous-specific education program.

The agency uses a five-point scale to score each program against the criterion — based on impact, strategic significance, risk profile and program expenditure. For example, the impact of the program for Aboriginal and Torres Strait Islander people could be given a score of five if it is a program aimed entirely at Aboriginal and Torres Strait Islander people, or one if it is expected to have less than 5 per cent Aboriginal and Torres Strait Islander clients.

Once a score has been allocated to each criterion, a weighted score is calculated for each program. The agency has decided to weight all criteria equally.

| <i>Program</i>  | <i>Impact of program</i> | <i>Strategic significance</i> | <i>Risk profile</i> | <i>Program expenditure</i> | <i>Weighted score</i> |
|---|--------------------------|-------------------------------|---------------------|----------------------------|-----------------------|
| Pilot study in a community  | 4                        | 5                             | 5                   | 2                          | <b>4</b>              |
| Mainstream education program  | 2                        | 4                             | 5                   | 5                          | <b>4</b>              |
| Established Aboriginal and Torres Strait Islander education program | 5                        | 5                             | 2                   | 2                          | <b>3.5</b>            |

The results of the prioritisation process show that the pilot study and the mainstream education program are considered the highest priorities for evaluation.

- The pilot study has only a relatively ‘small’ aggregate impact in terms of expenditure and population, but it is assessed as being high priority for evaluation because the potential risk of the program to a child’s wellbeing is significant and it is a particularly innovative program in a strategically-important policy area where the existing evidence base is poor.
- The mainstream education program has a relatively small proportion of Aboriginal and Torres Strait Islander participants, but has a large budget and is designed to affect school attainment, an outcome that has been identified as being of very high priority to Aboriginal and Torres Strait Islander people and the government.
- The established Aboriginal and Torres Strait Islander education program is ranked as a lower priority for evaluation because it has a relatively small budget and has already been subject to recent evaluation.

For transparency around how evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are prioritised, all Australian Government agencies should produce, on an annual basis, a rolling Three Year Evaluation Forward Work Plan, which details:

- policies and programs within their portfolio that contribute to government-wide evaluation priorities aimed at improving the lives of Aboriginal and Torres Strait Islander people



- 
- the methodology used to categorise the high priority policies and programs (based on the above criteria)
  - a plan for how/when over the next three years the agency's identified policies and programs will be evaluated (or how they will become ready for evaluation)
  - how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges were centred as part of the prioritisation process.

#### **Questions to consider when deciding what policies and programs to evaluate**

- What policies and programs administered by the agency are within scope of the Indigenous Evaluation Strategy?
- Have Aboriginal and Torres Strait Islander people been engaged in the process of deciding what policies and programs to evaluate?
- Does the agency have an evaluation work plan in place? Is it publicly available?

#### **Further reading on deciding what to evaluate**

*Indigenous Evaluation Strategy, Draft Background Paper*, chapter 6.

Department of Industry, Innovation and Science 2019, *Evaluation Strategy 2017 to 2021*, Canberra.

---

### **3 Evaluation planning, design and conduct**

Once a decision has been made to evaluate a policy or program, detailed evaluation planning should take place. This involves deciding on the evaluation questions, the type of evaluation, the approach and methods that will be used, the data needed to answer the evaluation questions, and who will conduct the evaluation. The evaluation plan should also identify key stakeholders, include a stakeholder engagement plan, a budget and an evaluation timeline.

This section outlines some of the key decisions that agencies will need to make during an evaluation. Some of these decisions will have to be made during the planning phase, others while the evaluation is being conducted, or later when the findings are released and acted upon (and some decisions will need to be revisited over the course of an evaluation). Knowing what key decisions need to be made before, during and after an evaluation phase can help agencies pre-empt problems and potentially address them during the planning phase.

#### **Deciding who will conduct the evaluation**

A decision about whether an evaluation will be conducted in-house by the agency or contracted out is usually made after some evaluation planning has taken place, but in most cases evaluators will be involved in developing a more detailed evaluation plan.

Independence of evaluators from policy and program management will enhance credibility, although close interaction between evaluators and policy and program staff can also be important for informing evaluation design. One option is for some parts of an evaluation to be undertaken externally (for example, data collection) and other parts undertaken internally by the agency. However, in all cases there should be functional independence between evaluators and those who make decisions about policies and program design and delivery.

The decision about whether the evaluation is undertaken internally or is externally commissioned should be based on who is best placed to undertake the evaluation given its purpose, circumstances and the type of evaluation. Agencies should ensure that evaluators (whether in-house or externally commissioned) have the requisite skills and experience to implement the Strategy's principles — including centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (noting that evaluations can use sophisticated methods that require particular skills and expertise).

There are important considerations when agencies commission external evaluators to ensure that contracting arrangements address the need for Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges to be centred during evaluation (box 4).

---

**Box 4****Centring Aboriginal and Torres Strait Islander people when commissioning evaluators**

While both internal and external evaluations should centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges, many decisions that an evaluation commissioner makes (such as those related to evaluation timelines, resourcing or scope) have the potential to affect how an evaluation is conducted. As such, commissioners should carefully consider how their commissioning practices may affect how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred. These include:

- ensuring that Aboriginal and Torres Strait Islander evaluators are able to bid for contracts by using open tender processes or procurement panels that include Aboriginal and Torres Strait Islander evaluators
- including an evaluator's experience in conducting research with Aboriginal and Torres Strait Islander people and their relationships with the communities involved as part of the selection criteria
- commissioners ensuring sufficient funding and time is allow for appropriate engagement with Aboriginal and Torres Strait Islander people during an evaluation, and for conducting the evaluation in line with ethical research practices (including allowing time for ethics clearance processes if required)
- ensuring that contracts and agency approval processes do not prevent findings being communicated back to Aboriginal and Torres Strait Islander participants and communities
- commissioners and evaluators working together to ensure ongoing communication with Aboriginal and Torres Strait Islander people involved with the evaluation.

---

## **Evaluation in practice: Good practice commissioning in a small-scale evaluation of a mainstream program**

A business coaching program is aimed at helping small to medium business enterprises expand. It is not targeted at Aboriginal and Torres Strait Islander businesses, and administrative records show that few Aboriginal and Torres Strait Islander businesses are using the program. Program managers want to investigate how to improve the program for Aboriginal and Torres Strait Islander, and other, users.

Aboriginal and Torres Strait Islander business experts were asked to provide feedback on the evaluation plan, which noted that Aboriginal and Torres Strait Islander businesses had difficulty accessing the program. Their feedback was incorporated into the evaluation plan. It was decided that a small case study might be the best way to identify challenges, while keeping evaluation costs low.

The agency decided to contract one evaluator to undertake the general evaluation and data analysis. A second evaluator, with demonstrated experience working with Aboriginal and Torres Strait Islander people, was contracted to conduct the case study. This allowed the agency to draw on specialised skills for the evaluation team.

The commissioner and evaluators had regular catch-ups to ensure that findings were integrated and shared during the evaluation. The data analysis found that while Aboriginal and Torres Strait Islander businesses did equally as well as non-Indigenous businesses from accessing the program, smaller Aboriginal and Torres Strait Islander businesses were less likely to apply for the program. The case study then explored these issues in detail with two different businesses.

The evaluation contracts stipulated that a results interpretation session would be conducted to ensure that findings from both studies were brought together. Those who gave feedback on the evaluation plan were invited to the session. Each participating business was given a short summary of the findings and their data.

---

### **Questions to consider when deciding who will conduct the evaluation**

- Will the evaluation be conducted by internal or external evaluators, or some combination of the two? Are potential evaluators objective and independent of those who make decisions about policy and program design and delivery?
- Are evaluators culturally and technically competent? Do potential evaluators have the necessary skills to conduct the evaluation with Aboriginal and Torres Strait Islander people effectively? Are evaluators Aboriginal and/or Torres Strait Islander people, or do they have existing partnerships with Aboriginal and Torres Strait Islander people?
- Who will be responsible for tendering, project management and quality assurance? Are the process and criteria used to make tender decisions transparent? Are evaluators selected through an open and transparent process?
- Have the terms of reference/request for tender been shaped by Aboriginal and Torres Strait Islander people and community input — for example, in terms of the required level of engagement, ethics requirements, evaluation outcomes, and treatment of intellectual property?
- Would engaging Aboriginal and Torres Strait Islander evaluators increase the relevance, cultural safety and credibility of the evaluation?
- Do procurement processes provide opportunities for Aboriginal and Torres Strait Islander evaluators to bid for contracts?

### **Determining the purpose and scope of the evaluation**

An important early step to a high-quality evaluation is being clear about its purpose and intended use (why the evaluation is being undertaken, who it is being undertaken for, and how it will be used). Answering these questions requires being clear about the policy or program's objectives and the levers that the policy will use to achieve its intended effects. The policy or program's logic model (section 1) will help evaluators identify the outcomes that need to be measured. If a program logic has not already been developed, this should be done as part of a detailed evaluation plan.

The evaluation's objective and questions should be developed with evaluation users and relevant stakeholders. The scope of an evaluation should consider the length and timing of the evaluation, available resources, target groups and geographical areas to be covered by the evaluation.

---

### **Questions to consider when determining the purpose and scope of the evaluation**

- Why is the evaluation being undertaken? Who is it being undertaken for, and how will the evaluation findings be used?
- Does the program logic clearly state how the program outcomes are expected to be achieved and how they are linked to higher level outcomes?
- What are the evaluation questions? Have Aboriginal and Torres Strait Islander people, and other affected by the policy or program, had an input into what the evaluation questions and evaluation approaches and methods should be?
- For mainstream programs, are the experiences of Aboriginal and Torres Strait Islander people given sufficient consideration?

### **Identifying and engaging with those affected by the policy or program**

Engagement with those affected by a policy or program should take place early and throughout the stages of an evaluation. Those affected by policy or program and potential evaluation users could include: government agencies, service providers, community organisations and policy or program recipients.

For evaluations conducted under the Strategy, agencies should engage with Aboriginal and Torres Strait Islander people and communities. Genuine engagement on evaluation questions, approaches, methods and reporting is a key way that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges can be centred in evaluations.

Engagement will look different depending on the policy or program being evaluated, the groups and communities affected and the scope and purpose of an evaluation. The stakeholder engagement plan will help agencies identify opportunities for, and barriers to, engagement. Opportunities for engagement with Aboriginal and Torres Strait Islander people during evaluation could include:

- engaging with Aboriginal and Torres Strait Islander peak bodies or community representatives during the evaluation planning phase
- Aboriginal and Torres Strait Islander leadership and/or representation on evaluation governance or steering committees
- engaging Aboriginal and Torres Strait Islander evaluators to lead evaluation planning and conduct
- governance arrangements for evaluations within an agency (for example, the agency's evaluation committee) including Aboriginal and Torres Strait Islander members.

---

## **Evaluation in practice: Building genuine engagement with Aboriginal and Torres Strait Islander people into evaluation design**

To build engagement with Aboriginal and Torres Strait Islander people into the design of an evaluation of an Indigenous-specific early childhood program, an evaluation plan was developed as a separate project.

External evaluators with experience and expertise working with Aboriginal and Torres Strait Islander people were commissioned to develop a detailed evaluation plan for the evaluation. The contract stipulated that engagement with Aboriginal and Torres Strait Islander people was to be a key deliverable for the project and that Aboriginal and Torres Strait Islander people were to be engaged in the planning phase of the evaluation, and throughout the course of the evaluation.

As part of the evaluation plan project, a steering committee with majority Aboriginal and Torres Strait Islander members was set up. The steering committee worked with the evaluators to develop the evaluation questions, choose methods and suitable indicators, provide guidance on ethics and develop a stakeholder engagement plan. The steering committee continued to guide the evaluation during the conduct phase.

## **Questions to consider when identifying and engaging with those affected by the policy or program**

- Who will be affected by the policy or program?
- Has sufficient time been allocated for engagement, including relationship building and establishing trust with those affected by the policy or program?
- Have opportunities for engagement been maximised, such as by providing those affected by the policy or program with information about the purpose and conduct of the evaluation opportunities and channels to participate? Have any potential barriers to engagement been identified and mitigated?
- Have Aboriginal and Torres Strait Islander stakeholders with an interest in the evaluation been identified? Can Aboriginal and Torres Strait Islander people participate in the evaluation on an equal footing with others? Are there language barriers or other factors that might limit Aboriginal and Torres Strait Islander people participating? Have cultural norms and standards for engaging with Aboriginal and Torres Strait Islander communities been adopted?
- Have Aboriginal and Torres Strait Islander people had input into how engagement takes place? Do engagement methods take account of contextual factors? Has there been consideration about whether Aboriginal and Torres Strait Islander languages should be used?
- What governance arrangements will need to be put in place? Are Aboriginal and Torres Strait Islander people represented in evaluation governance arrangements?
- What mechanisms are in place to make evaluation processes transparent?

## Choosing an appropriate evaluation type, approach and method

Evaluations are initiated for a range of reasons, and they can answer a broad range of questions, but there are three main questions they seek to answer.

- How well is the policy or program being delivered?
- What difference did the policy or program make?
- Do the benefits of the policy or program justify the costs?

There are also three broad types of evaluation — process, impact and value-for-money. The evaluation type and approach adopted will depend on the questions that need to be answered, the kind of policy or program being evaluated, the circumstances under which the policy or program is implemented and the resources available for the evaluation (table 2).

**Table 2      What type of evaluation? It will depend on the purpose and the questions you want answered**

| <i>Evaluation type</i>  | <i>Questions</i>   | <i>Purpose</i>  |
|---|--|---|
| <b>Formative or process evaluations</b> — these evaluations are undertaken early in the development or implementation of a policy or program. | How is the policy or program delivered?<br>Is the program being delivered as intended?<br>Is the policy or program appropriately targeted?<br>How effective has implementation been so far?<br>What are the policy or program's strengths and weaknesses?<br>Can the policy or program be improved to achieve better outcomes? | These evaluations help us to better understand the mechanisms at play in successful and less successful policies and programs. They can help shape a policy or program to perform better. The evaluation can assist in improving an initiative as it is rolled out and can provide a baseline for future evaluations. |
| <b>Summative, outcome or impact evaluations</b> — these evaluations judge the overall merit, worth and impact of a policy or program.         | What difference did the policy or program make?<br>Has the policy or program achieved its objectives? Does the evidence support the theory?<br>Has the policy improved outcomes? If so, by how much?<br>Did the policy affect groups of users differently?   | These evaluations are undertaken for lesson-learning (they can be used to inform decisions about whether to expand, cease, replicate or scale up a program) and accountability.   |
| <b>Economic evaluations</b> — these evaluations quantify the net benefit of a policy or program.  | Do the benefits justify the costs, or was it worth it?<br>Are there alternative approaches that would result in lower costs for the same benefits?   | These evaluations quantify the value of policies and programs and can be used to compare options. They are undertaken for accountability and resource allocation decisions.   |



---

There are number of questions agencies should ask when deciding on a type of evaluation.

- How complex is the relationship between the policy or program and the outcome (simple relationships can be investigated using process evaluations)?
- How important is it to control for other factors affecting the outcome (if it is important to control for other factors this will point to an impact evaluation)?
- How significant are the potential outcomes of the policy or program in terms of its contribution to an overall policy objective?
- How significant is the policy or program in terms of resources?
- How is the policy or program being implemented (will it allow for a counterfactual)?<sup>1</sup>

The choice of type of evaluation and approach should maximise the rigour and credibility of the evaluation given the questions the evaluation seeks to answer, the nature and scale of the policy or program, resourcing and data constraints. Answering the evaluation questions will often involve a combination of evaluations over different stages of implementation. For example, a process or formative evaluation could be undertaken in the early stages of implementation for a new policy or program to ensure that it is being delivered as intended, with an impact evaluation undertaken after the policy or program has been operating for several years.

Rigour is required for all types of evaluations and requires:

- ensuring findings about cause and effect are well founded
- being clear about the populations, settings or circumstances to which evaluation results can be generalised
- that the indicators or metrics used accurately capture the outcomes of interest.

Both quantitative and qualitative data have a place in evaluation and together can tell a more complete story. Quantitative data can be used to track trends over time and understand patterns and differences in outcomes between policy and program participants. Qualitative methods can provide in-depth insights into people's experiences and perspectives. They have a particular strength in addressing questions of why and how. Qualitative methods are also sometimes preferred in Aboriginal and Torres Strait Islander contexts because they allow participants to express themselves using their own words and concepts. They can also help counteract power differences by giving voices to those with less power. But on their own, qualitative methods are typically not sufficient to measure policy or program outcomes because they rarely measure whether real change has occurred.

Employing triangulation across methods and sources of data (including combining quantitative and qualitative methods) can make each method and data source more valuable and strengthen validity. Surveys, for example, could ask a mix of scale questions (quantitative) and opened-ended questions (qualitative). Evaluators could also ask different

---

<sup>1</sup> Adapted from HM Treasury 2011, *The Magenta Book: Guidance for Evaluation*, p. 24.

---

people about the same issue (for example, an evaluation of teaching methods could get feedback from students, teachers and parents).

Indigenous evaluation approaches place an Indigenous lens over the policy or program being evaluated, as well the evaluation process itself, to ensure that the values, experiences, knowledges and interests of Aboriginal and Torres Strait Islander people are recognised and upheld. Culturally valid methods, such as yarning (storytelling), ganma (knowledge sharing) and dadirri (listening) can be used to ensure Aboriginal and Torres Strait Islander knowledges, perspectives and world views are embedded in evaluations.

For evaluations of mainstream policies and programs, the approach will depend on the type of policy or program being evaluated, the number or proportion of Aboriginal and Torres Strait Islander people affected by the policy or program and what is already known about the impacts of the policy or program on Aboriginal and Torres Strait Islander people. Options for assessing the impact of mainstream policies and programs on Aboriginal and Torres Strait Islander people include:

- examining policy or program outcomes for Aboriginal and Torres Strait Islander people and non-Indigenous people to determine if there are differential impacts
- using focus groups or case studies to examine particular issues affecting Aboriginal and Torres Strait Islander people more closely
- undertaking a targeted evaluation that looks specifically at the impact of a mainstream policy or program on Aboriginal and Torres Strait Islander people.

Chapter 4 of the *Indigenous Evaluation Strategy, Draft Background Paper* contains further information on the types of evaluations and when each type may be appropriate.

---

### **Evaluation in practice: Evaluating to understand why there are poorer outcomes for Aboriginal and Torres Strait Islander people**

An evaluation of a large mainstream employment program found that Aboriginal and Torres Strait Islander people had poorer employment outcomes than non-Indigenous people when participating in the program. The agency responsible for delivering the program wanted to understand why this was the case and planned a further evaluation focusing specifically on outcomes for Aboriginal and Torres Strait Islander people.

As a first step, the evaluation used administrative data from program participants to examine outcomes for different groups of Aboriginal and Torres Strait Islander people. This analysis found that some of the difference in employment outcomes between Aboriginal and Torres Strait Islander participants and non-Indigenous participants was because of higher caring burdens — carers typically had poorer outcomes from the program regardless of their Indigenous status. This led to a recommendation that further support be provided for all carers so that they could participate fully in the program.

The quantitative analysis also found that Aboriginal and Torres Strait Islander people had disproportionately poor employment outcomes in some regions. The evaluation adopted a case study approach, choosing several regions where outcomes were particularly good and particularly poor to study in more detail. Focus groups were held with service providers and participants in selected regions to better understand regional differences and to collect examples of good practice to share across regions. Surveys were also conducted with participants in case study regions to supplement administrative data on employment outcomes. Data collected were used to explore possible unintended consequences (both positive and negative) of the program.

### **Questions to consider when choosing an appropriate evaluation type, approach and method**

- What evaluation types, approaches and methods will best answer the evaluation questions? Should the evaluation be a formative, process, impact evaluation (or some combination)? Is an economic evaluation required? Is there merit in using different approaches for each evaluation question?
- What qualitative and quantitative methods should be used?
- What evaluation methods are most appropriate for Aboriginal and Torres Strait Islander participants? Do the methods used draw on Aboriginal and Torres Strait Islander expertise, perspectives, methodologies and knowledges?

## **Identifying and collecting data**

The questions, scope and approach of an evaluation will determine what data are needed for an evaluation. A variety of different data sources can be used for evaluation (box 5).

---

## Box 5 Sources of data

There are many different sources of data to consider when seeking to collect data for an evaluation. These include:

- *Administrative data* — data that are collected as part of administering a policy or program. They typically include registrations, transactions and general record keeping. Administrative sources, however, only include data on individuals that interact with a policy or program. For example, they can identify who is receiving unemployment benefits, but not what happens to them when they leave benefits. This weakness may be addressed through data linkage where multiple datasets are linked together.
- *Monitoring data* — monitoring data are performance management data that are collected as a policy or program is delivered. It generally covers all aspects of a policy or program including inputs, processes, outputs and outcomes. A particularly useful feature of monitoring data is that they provide data throughout the life of a policy. However, care needs to be taken to ensure they are of high quality. If data collection is seen as not useful or an administrative burden, there is little incentive to collect high quality data.
- *Survey data* — there are large-scale surveys that potentially could be used for evaluation, including the Census of Population and Housing or the Longitudinal Study of Indigenous Children. Using existing survey data will be less expensive and will place less burden on participants than collecting new data, but the timing, representativeness or specific questions asked can limit their usefulness for evaluation. Surveys can also be used to collect data as part of evaluation or monitoring.
- *Qualitative data* — qualitative data will most likely have to be collected as new data, although some monitoring data are qualitative. A variety of methods may be employed including interviews, focus groups, case studies and observations.

Source: adapted from HM Treasury 2020, *The Magenta Book: Central Government Guidance on Evaluation*.

Data for evaluation should be fit-for-purpose and the costs of collecting data for evaluation balanced with the benefits.

As discussed earlier, it is important to plan early for data collection. It allows data requirements to be built into a policy or program's delivery, for baseline and counterfactual data to be collected, and for data approvals to be obtained.

Data should be reflective of the lives and experiences of Aboriginal and Torres Strait Islander people. To improve the relevance of data to Aboriginal and Torres Strait Islander people, and improve trust in its collection and quality, Aboriginal and Torres Strait Islander people should be involved in data planning, collection and governance during evaluation.

Given the costs of collecting new data, existing data should be used wherever possible. However, most evaluations will involve at least some new data collection. Where Aboriginal and Torres Strait Islander people are a small percentage of participants in the policy or program, oversampling may be necessary to collect enough responses for robust results.

---

## Evaluation in practice: Collecting baseline data

An evaluation of a literacy program in primary schools intends to compare outcomes before and after the program has been implemented to determine its impact. An evaluation plan including a data plan was developed before the program was implemented in schools. The program logic helped the agency to identify the key outcomes of interest, as well as indicators that would be useful to track progress and provide context for a future evaluation.

Some of these indicators were already collected through regular school testing surveys conducted annually. When rolling out the program in schools, the agency ensured that participants were asked for consent to use data from their regular school testing survey in the evaluation. Participants also completed a survey prior to starting the program to collect baseline data on other key indicators that the agency expected the program to influence but that were not included in the regular survey.

When an outcome evaluation is conducted several years later, the evaluators are able to access the earlier data to use as a baseline. The agency had clearly documented the data sources and collection methods so that the same methods can be replicated to collect data from participants after they had participated in the program.

## Evaluation in practice: Oversampling to improve data quality

An evaluation of a mainstream health program wants to know about outcomes for Aboriginal and Torres Strait Islander participants, who make up around 5 per cent of all program participants. Data on outcomes will be collected via interviews with participants when they are discharged from hospital, and follow up surveys sent to participants six months later. In previous evaluations of similar programs, evaluators found that response rates for Aboriginal and Torres Strait Islander participants were lower than those for non-Indigenous participants, particularly for the follow up survey. However, response rates were higher for Aboriginal and Torres Strait Islander participants who were referred to the program by an Aboriginal community-controlled health service than those referred by a mainstream health service.

To ensure there are sufficient survey responses from Aboriginal and Torres Strait Islander participants, evaluators plan to *oversample* — this means including more Aboriginal and Torres Strait Islander people in the interviews and surveys than would be the case if respondents were chosen randomly. The number of Aboriginal and Torres Strait Islander people targeted for interviews will be decided based on response rates to previous surveys, and the need to be able to break down the sample of Aboriginal and Torres Strait Islander people by key demographic characteristics (including gender, age and geographical location).

To encourage higher response rates, evaluators will work with Aboriginal community controlled health services to design the interview and survey questions, and provide additional training to interviewers on culturally safe data collection techniques. More information will be provided to participants on the benefits of responding to the survey, including being able to tailor services to better meet their needs. The evaluation team will also produce summary reports of program statistics by region so that service providers and community organisations can see how their region compares with others.

---

### Questions to consider when identifying and collecting data

- What data are needed to answer the evaluation questions? What data are needed to produce credible results? Can multiple sources of data be used to improve the rigour of results?
- Do outcome measures align with the intended objectives of the policy or program? Do outcome measures reflect what's important to Aboriginal and Torres Strait Islander people's lives?
- Are data needed on policy or program inputs, the context in which the policy or program is being delivered or any unintended consequences of the policy or program?
- When should data on outcomes be collected?
- What existing data are available to be used? Is approval needed to use them?
- What additional data should be collected? How will data be collected? Do the data collection methods consider the needs, rights and cultural safety of respondents? Are the benefits from collecting additional data likely to exceed the costs of doing so?
- Are collected data of high quality? What are the limitations of the data? How can these be overcome?
- Have Aboriginal and Torres Strait Islander people been engaged in planning, collecting and managing data used in the evaluation?
- Are there clear processes in place for access to and release of data?

### Ensuring ethical and culturally safe evaluation

Ethics in evaluation is concerned with good conduct — the conduct of evaluators, the integrity of the evaluation process, and the protection of participants. It also requires that Aboriginal and Torres Strait Islander peoples values and cultural practices are observed during evaluations involving them and their communities.

Evaluations should be conducted ethically, although not all require formal ethics review (box 6). However, all evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people undertaken by Australian Government agencies should be subject to a systematic assessment of ethical risks, which should be included in evaluation reports. Ethical conduct during the evaluation should also be clearly documented.

Evaluations conducted under the Strategy should follow established guidelines for research with Aboriginal and Torres Strait Islander people, such as those produced by the Australian Institute for Aboriginal and Torres Strait Islander Studies (AIATSIS) and the National Health and Medical Research Council (NHMRC) listed in the further reading section.

Agencies should ensure that evaluation budgets and timeframes are sufficient for evaluators to fulfil ethical requirements. This means allowing enough time for ethics risk assessment and formal ethical review, if required, as well as for evaluators to undertake meaningful engagement with Aboriginal and Torres Strait Islander people, build capability and report back to evaluation participants.

---

## Box 6      **When is formal ethical review necessary?**

Formal review by an ethics committee will be required for some but not all evaluations conducted under the Indigenous Evaluation Strategy. In some cases where evaluation relies exclusively on existing or de-identified data, formal ethical review may not be necessary. However, agencies and evaluators should still be guided by ethical practices in these cases and research quality may be improved by getting advice from an ethics committee with experience in Aboriginal and Torres Strait Islander research.

All evaluation projects should include a systematic assessment of ethical risks. Decisions on whether to request formal review by an ethics committee should be clearly documented. Ultimate responsibility for deciding whether formal ethical review is required should lie with agencies rather than with commissioned evaluators. However, agencies should take advice from commissioned evaluators when undertaking ethical risk assessment.

Where formal ethics review is required, it should be undertaken by an ethics committee with expertise in Aboriginal and Torres Strait Islander research and guided by recognised ethical guidelines for research involving Aboriginal and Torres Strait Islander people.

### **Questions to consider about ensuring ethical and culturally safe evaluation**

- Has the evaluation design and method been critiqued against Aboriginal and Torres Strait Islander ethics frameworks for research and evaluation?
- Has there been a systematic assessment of ethical risks associated with the evaluation? Is formal review by an ethics committee required?
- If ethics review is required, does the ethics review body have expertise appropriate for evaluation involving Aboriginal and Torres Strait Islander people?
- Have sufficient time and resources been allowed for evaluators to meet ethical requirements?
- In what ways can the evaluation methods and processes be adapted to adopt ethical principles?
- How are participants' contributions of information, knowledges and time being respectfully and appropriately recognised and valued by evaluators and evaluation commissioners?
- How will the intellectual property from the evaluation be treated? Has this been negotiated with communities involved?
- How will consent be obtained from participants?
- Are there any confidentiality or privacy concerns? How will they be addressed?
- Are there any ethics or research protocols that need to be followed in the communities involved in the evaluation?

## **Interpreting and verifying the results**

Interpreting results in an evaluation involves assessing the evidence (and its limitations) to give the evidence meaning and answer the evaluation questions. Ideally, during the planning

---

stage when considering what approach and data to use, thought should be given to how the evidence will be used to make an assessment.

The interpretation of evidence should be systematic, logical and clearly articulated. Any limitations should be acknowledged with specific reference to how they may affect the validity and generalisability of findings. Alternative explanations for observed outcomes should be explored.

Interpreting results with Aboriginal and Torres Strait Islander people allows for interpretations and conclusions to incorporate Aboriginal and Torres Strait Islander knowledges and perspectives. This improves the validity and usefulness of findings for policy and program designers and service providers, as well as for Aboriginal and Torres Strait Islander evaluation users. When interpreting the results, there may not always be consensus, and where there are differing views they should be reflected in the evaluation report.

### **Evaluation in practice: Workshopping evaluation findings with participants**

An evaluation of a health program operating in urban areas collected quantitative and qualitative data via surveys, focus groups with program participants, and interviews with service providers.

Evaluators held workshops with service providers and evaluation participants to present preliminary results and to seek clarification and feedback on the key findings emerging from the evaluation. The workshop participants provided further context on the preliminary findings, which allowed the evaluators to better understand the reasons behind some conflicting results. The evaluators received useful feedback on what was important to participants, allowing the final evaluation report to be tailored to users' needs. The workshops also allowed service providers to get timely access to information that they could use to improve program delivery.

Whether an evaluation is conducted internally or using external evaluators, a key part of managing an evaluation is ensuring quality throughout the process. Ongoing quality assurance measures (such as peer review and steering committees) can help ensure that evaluation design, planning and delivery are properly conducted and conform to professional, ethical and analytical standards. Peer review, for example, allows experts independent of the policy or program and the evaluation to assess whether the evaluation questions, design and conduct are fit-for-purpose and that ethical requirements have been met. They can be conducted internally (by someone not connected to the policy or evaluation) or externally, by someone who is an expert in the subject area or in the evaluation approach or method used.



---

### Questions to consider when interpreting and verifying the results

- Do the findings answer the evaluations questions? Are the findings supported by data and evidence? What limitations have been identified? Do they call into question the validity of the findings?
- Have alternative explanations been explored? Has any disagreement been acknowledged?
- Have Aboriginal and Torres Strait Islander people had opportunities to interpret the findings or provided feedback on the findings? Do Aboriginal and Torres Strait Islander people have a similar interpretation of the evidence and data to the evaluator?
- Are the data collection methods, assumptions and value judgements clearly documented and any limitations noted?
- How well might the findings apply to other policies, programs or contexts?
- What quality assurance processes are in place?

### Further reading on evaluation planning, design and conduct

*Indigenous Evaluation Strategy, Draft Background Paper*, chapters 4, 5, 7, 8 and 9.

Australian Evaluation Society nd, 'Resources', <https://www.aes.asn.au/resources.html>

Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) 2012, *Guidelines for Ethical Research in Australian Indigenous Studies*.

—— 2019, *The AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*, Consultation draft.

Better Evaluation nd, 'Evaluation practice in Aboriginal and Torres Strait Islander settings', <https://www.betterevaluation.org/en/themes/evaluation-practice-aboriginal-and-torres-strait-islander-settings>.

Better Evaluation nd, 'Manager's guide to evaluation', [https://www.betterevaluation.org/managers\\_guide](https://www.betterevaluation.org/managers_guide).

Centre for Epidemiology and Evidence 2019, *Commissioning Economic Evaluations: A Guide*, Evidence and Evaluation Guidance Series. Population and Public Health Division, NSW Ministry of Health, Sydney.

Department of the Prime Minister and Cabinet 2018, *Indigenous Advancement Strategy Evaluation Framework*.

HM Treasury 2012, *Quality in Policy Impact Evaluation: Understanding the Effects of Policy from other Influences (Supplementary Magenta Book Guidance)*.

—— 2012, *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence (Supplementary Magenta Book Guidance)*.

—— 2020, *The Magenta Book: Central Government Guidance on Evaluation*.

---

Hunt, J., 2013. *Engaging with Indigenous Australia — Exploring the Conditions for Effective Relationships with Aboriginal and Torres Strait Islander Communities*, Issues paper 5, Closing The Gap Clearinghouse.

Kelaher, M., Luke, J., Ferdinand, A., Chamravi, D., Ewen, S. and Paradies, Y. 2018, *An Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health*, The Lowitja Institute, Melbourne.

National Health and Medical Research Council (NHMRC) 2018, *Keeping Research on Track II*.

——— 2020, *Ethical Guidelines for Research with Aboriginal and Torres Strait Islander Peoples*.

Rogers, P., Hawkins, A., McDonald, B., Macfarlan, A., Milne, C. 2015, *Choosing appropriate designs and methods for impact evaluation*, Department of Industry, Innovation and Science, Canberra.

World Health Organisation 2013, *WHO Evaluation Practice Handbook*, Geneva.

---

## 4 Reporting evaluation findings

The aim of all evaluations conducted under the Strategy should be to influence decisions about policy and program design and implementation. To do so, evaluation findings should be shared publicly in forms that all evaluation users — including Aboriginal and Torres Strait Islander people — find useful. Transparency increases evaluation quality, accountability and opportunities for learning from findings.

To be useful, evaluation reports should be logical, clear and complete. That is, evaluation reports should:

- explain the policy or program being evaluated and the evaluation questions
- outline the methodology and data used
- present evidence to support findings
- clearly articulate limitations.

All evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people conducted under the Strategy should include a short plain English summary of the evaluation report (box 7).

Evaluations of mainstream policies and programs under the Strategy should report on the impacts and effectiveness of the policy or program for Aboriginal and Torres Strait Islander people.

### Box 7 Plain English summaries of evaluation reports

To make evaluation findings accessible to all evaluation users, evaluation reports should include a short plain English summary, including details such as the:

- main findings of the evaluation (and depending on timing, agency responses to the main findings)
- methodology and data used
- extent of engagement with Aboriginal and Torres Strait Islander people, organisations and communities and others affected by the policy or program during the evaluation
- whether a formal ethical review was undertaken and measures taken to ensure that the evaluation meets ethical standards
- limitations of the evaluation
- time and cost to complete the evaluation.

All evaluation reports should be published on agency websites, shared on the new online Indigenous Evaluation Clearinghouse and disseminated to evaluation users. Where evaluation reports cannot be made public (such as in cases where there are concerns about privacy or culturally sensitive information) the reasons for not publishing should be included in a published summary.

---

Evaluation results are more likely to be used if they are synthesised and tailored for specific users. Agencies should consider ways to disseminate evaluation findings to evaluation users (including verbal feedback or information sessions, and summaries in languages other than English if appropriate). A knowledge dissemination plan should be agreed on with Aboriginal and Torres Strait Islander people, organisations and communities at the evaluation planning stage.

### **Evaluation in practice: Presenting evaluation findings for different audiences**

A mainstream education program aimed at preventing school dropout among secondary school students is evaluated to measure its impact and examine how effectively it has been implemented in schools. As well as producing an evaluation report that documents the evaluation methods, data, findings and limitations in detail, the agency commissions the evaluators to produce a range of fact sheets targeted at different audiences: parents, teachers and students. The fact sheets highlight examples of good practices that had been identified during the evaluation as contributing to better outcomes.

The findings of the evaluation are presented to staff and managers in the agency at a seminar. The seminar focuses on both the detailed findings of the evaluation for the implementation and design of the program being evaluated, as well as broader lessons that may be relevant to other program areas.

The evaluation was part of a broader program of research commissioned and conducted by the agency. The evaluation findings are incorporated into a synthesis report on factors affecting school retention that is published on the agency's online portal for sharing evidence and presented at conferences attended by teachers and education administrators.

---

### Questions to consider when reporting evaluation findings

- Are evaluation methods and data described in detail in the report?
- Are the limitations of the evaluation analysis, data and results clearly documented in the report?
- Is the basis for the evaluation findings transparent?
- Is ethical conduct during the evaluation clearly documented?
- Are participants' contribution recognised in evaluation reports?
- Are any conflicts of interest, and how such conflicts were managed during the evaluation, disclosed?
- Are differential effects observed across different populations? Does the evaluation report discuss the impacts and effectiveness of the policy or program for Aboriginal and Torres Strait Islander people?
- Is there a plain English summary of the evaluation findings?
- How will the findings be shared and disseminated? Are reports published and easy to find?
- Is appropriate attention given to addressing concerns about privacy, confidentiality and culturally sensitive information?
- Will it be necessary to present the findings in different formats for different users?
- Have results been communicated back to the Aboriginal and Torres Strait Islander communities or organisations engaged in the evaluation?
- Are there mechanisms in place for knowledge translation? Have results been translated/interpreted in a culturally-sensitive manner?
- Has the evaluation report been submitted to the Indigenous Evaluation Clearinghouse?

### Further reading on reporting evaluation findings

*Indigenous Evaluation Strategy, Draft Background Paper*, chapter 7.

National Health and Medical Research Commission (NHMRC) 2018, *Keeping Research on Track II*.

---

## 5 Building capability and a culture of evaluation

The Strategy will be most effective in an environment where agencies want to know how their policies and programs are performing, and are prepared to experiment, share learnings and use evaluation results in policy making. Building evaluation capability — including cultural capability — is also important for producing high-quality evaluations with credible results.

In practice, this means that evaluative thinking — a form of critical thinking where evaluative questions are asked as a matter of habit — is embedded within agencies. Agency staff need to be able to plan for and manage evaluations, interpret and implement findings. They may also need to be able to commission external evaluators, including selecting an evaluator, managing an evaluation and assessing its quality.

### **Evaluation in practice: Building evaluation capability in a small agency**

A small agency wants to build their evaluation capability and ensure that staff have the skills to implement the Indigenous Evaluation Strategy. Given the size of the agency, it is unlikely to conduct evaluations internally, so capability building efforts will focus on improving skills at managing and commissioning evaluations.

The agency contracts an external evaluator to develop templates and checklists for program managers to use when planning evaluation and provides training on evaluation to managers and staff. Training uses a coaching approach where program managers are encouraged to bring their own program logic or evaluation plan to work to the session and other program managers are invited to provide feedback.

The agency works on building cultural capability, first by requiring staff involved in commissioning evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people to complete the Australian Institute for Aboriginal and Torres Strait Islander Studies' Core Cultural Learning foundation course. The agency also engages an Aboriginal and Torres Strait Islander organisation to provide coaching to staff to ensure that evaluation plans and program logic are culturally appropriate.

Program managers and staff involved in evaluation are also encouraged to seek out other opportunities to develop their evaluation or cultural capability skills by engaging with communities of practice, professional organisations, Aboriginal and Torres Strait Islander community controlled organisations and other evaluators within the Australian Public Service.

Transparency is vital for improving agencies' evaluation culture. As well as publishing evaluation reports, other documents about evaluation policies — including frameworks, strategies and forward work plans — should be published on agency websites so it is easy for stakeholders to see agencies' commitments on evaluation.

Embedding evaluation in the policy cycle is key to building evaluation culture and capability. Adopting measures discussed earlier in this guide — such as building on existing evidence when designing new policies and programs, and planning early for evaluation — can improve the relevance, credibility and usefulness of evaluation to decision makers.

---

But it is also critical that evaluation findings feed upward to decision makers and translate to action — be it the establishment of a new or reformed initiative, the cessation of a policy or program, or (in most cases) incremental change in light of the lessons learned. Evaluation findings should also inform the strategic priority setting of government, enabling decisions to be made about the best and most effective allocation of resources.

To maximise transparency and accountability, agencies should publish their management response to published evaluation reports (this could be included as part of the evaluation report, or published separately). This response should include an explanation about what they have learned, what they have changed in response to the findings, and any further action they intend to take.

### **Questions to consider when building capability and a culture of evaluation**

- What is the agency's approach to evaluation? Is evaluative thinking valued?
- Does the agency have an evaluation policy or framework? Is this publicly available? Does the agency have a central unit to plan and share lessons from evaluation?
- Are there processes that allow the findings of evaluations to be incorporated in decisions about policy or program changes? How will evaluation findings be shared within the agency?
- Does the agency have processes in place to build the capability to use and respond to evaluation findings?
- Do evaluation teams have the capacity, cultural and technical capability and experience to deliver high-quality evaluation?
- Does the agency provide opportunities or encourage staff to pursue opportunities to strengthen their technical and cultural capability? Do evaluation teams have the capability to incorporate Aboriginal and Torres Strait Islander knowledges into their evaluative thinking, including by asking questions that matter for Aboriginal and Torres Strait Islander people and collecting and using evidence in culturally safe ways?
- Does the agency have processes that seek to build capability among Aboriginal and Torres Strait Islander evaluators, organisations and communities?
- Is the agency open to review and feedback from Aboriginal and Torres Strait Islander people and organisations about its evaluation practices?
- Is it clear what the findings will be used for and what decisions they will feed into?
- Are responses to the findings of the evaluation published?

### **Further reading on building capability and a culture of evaluation**

*Indigenous Evaluation Strategy, Draft Background Paper*, chapters 7 and 8.

Better Evaluation nd, 'Manager's guide to evaluation', [https://www.betterevaluation.org/managers\\_guide](https://www.betterevaluation.org/managers_guide).



[www.pc.gov.au/ies](http://www.pc.gov.au/ies)