



Australian Government
Productivity Commission

Indigenous Evaluation Strategy

Productivity Commission Draft Background Paper

May 2020



This is a draft paper prepared for further public consultation and input. The Commission will finalise its report after these processes have taken place.

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, Indigenous Evaluation Strategy, Draft Background Paper*.

If you have adapted, modified or transformed this work in anyway, please use the following, *Source: based on Productivity Commission data, Indigenous Evaluation Strategy, Draft Background Paper*.

An appropriate reference for this publication is:

Productivity Commission 2020, *Indigenous Evaluation Strategy, Draft Background Paper*, Canberra, May.

Publications enquiries

Media, Publications and Web, phone: (03) 9653 2244 or email: 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).

Opportunity for further comment

You are invited to examine this draft background paper together with the *Indigenous Evaluation Strategy* and *A Guide to Evaluation under the Indigenous Evaluation Strategy* and provide comments by written submission to the Productivity Commission, preferably in electronic format, by 3 August 2020.

Further information on how to provide a submission is included on the study website: www.pc.gov.au/inquiries/current/indigenous-evaluation.

The Commission will prepare a final background paper and final *Indigenous Evaluation Strategy* after further submissions have been received and we have engaged with stakeholders. The Commission will provide the final background paper and Strategy to the Australian Government in October 2020.

Commissioners

For the purposes of this study the Commissioner is:

Romlie Mokak

Letter of direction

I, Josh Frydenberg, Treasurer, pursuant to Parts 2 and 4 of the *Productivity Commission Act 1998* hereby request the Productivity Commission to develop a whole-of-government evaluation strategy for policies and programs affecting Indigenous Australians. The Commission will also review the performance of agencies against the strategy over time, focusing on potential improvements and on lessons that may have broader application for all governments.

Background

A number of high profile reports have highlighted the need for more evaluation of policies and programs that have an impact on Indigenous Australians. For example, the Commission's *Overcoming Indigenous Disadvantage Report 2016* found that only a relatively small number of programs have been rigorously evaluated.

Improving outcomes for Indigenous Australians depends on agencies with responsibility for policies and programs affecting Indigenous Australians undertaking meaningful evaluations. The Commission is to develop a strategy to guide that evaluation effort.

Scope

The Commission should develop an evaluation strategy for policies and programs affecting Indigenous Australians, to be utilised by all Australian Government agencies. As part of the strategy, the Commission should:

- establish a principles based framework for the evaluation of policies and programs affecting Indigenous Australians
- identify priorities for evaluation
- set out its approach for reviewing agencies' conduct of evaluations against the strategy.

In developing the strategy, the Commission should consider:

- how to engage Indigenous communities and incorporate Indigenous knowledge and perspectives
- ethical approaches to evaluations

-
- evaluation experience in Australia and overseas
 - relevant current or recent reviews commissioned or undertaken by Australian, state, territory or local government agencies
 - the availability and use of existing data, and the further development of other required data and information
 - areas in which there may be value in the Productivity Commission undertaking evaluation
 - how to translate evidence into practice and to embed evaluation in policy and program delivery.

Process

The Commission should consult widely on the strategy, in particular with Indigenous people, communities and organisations (such as the Empowered Community regions), and with all levels of government. It should also consult with non-Indigenous organisations, and individuals responsible for administering and delivering relevant policies and programs.

The Commission should adopt a variety of consultation methods including seeking public submissions.

The Commission should provide the evaluation strategy and forward work program to Government within 15 months of commencement.

The Hon Josh Frydenberg MP
Treasurer

[Received 10 April 2019]

Contents

Opportunity for further comment	iii
Letter of direction	iv
Abbreviations	x
Overview	3
1 What we have been asked to do and our approach	4
2 The current state of evaluation	5
3 A new approach — an Indigenous Evaluation Strategy	11
4 Review and future governance arrangements	37
Information requests	41
1 About the Strategy	45
1.1 Some background to the Strategy	45
1.2 What is evaluation and why evaluate?	47
1.3 Why a whole-of-government Indigenous evaluation strategy?	48
1.4 How we approached the task	53
1.5 Improving lives is the objective	57
1.6 A roadmap for the rest of the report	69
2 Evaluation in the APS — history and current practice	71
2.1 The ebbs and flow of evaluation history	71
2.2 The current and changing landscape of evaluation	75
2.3 How do Australian Government agencies currently approach evaluation?	81
3 The state of evaluation for policies affecting Indigenous people	95
3.1 What we heard about the current state of evaluation	96
3.2 What policies and programs have been evaluated?	100

3.3	What does the evidence tell us about what works (or does not work)?	111
4	Types of evaluation	121
4.1	Evaluation types, approaches and methods	122
4.2	How do you decide on an evaluation type, approach and method?	139
5	Evaluating for quality results	153
5.1	What makes for a high quality evaluation?	154
5.2	Engaging effectively with Aboriginal and Torres Strait Islander people and their knowledges	170
5.3	Evaluating mainstream policies and programs	181
6	What to evaluate	189
6.1	Why set evaluation priorities?	190
6.2	Approaches to setting evaluation priorities	192
6.3	Priority setting under the Indigenous Evaluation Strategy	198
6.4	Initial implementation and a process of review	212
7	Embedding a culture of evaluation and learning	215
7.1	What is a culture of evaluation and learning?	217
7.2	Enhancing the use of evaluations across the APS	219
7.3	Creating incentives to produce high-quality and useful evaluations	234
8	Evaluation capability	253
8.1	What is the current state of evaluation capability in the APS?	254
8.2	The skills required to undertake and commission evaluations	256
8.3	Capability to commission high quality evaluations	268
9	Data for evaluation	279
9.1	How are data currently used in evaluations?	280
9.2	Some issues raised about data for evaluation	283
9.3	Improving data for evaluation	295

10	Monitoring and reviewing the Strategy	319
10.1	Monitoring arrangements for the Strategy	320
10.2	Review arrangements for the Strategy	331
11	Bringing it all together: an Indigenous Evaluation Strategy	335
11.1	The key components of the Strategy	336
11.2	What principles should underpin the Strategy?	342
A	Conduct of the project	355
B	Information request to Australian Government agencies	(online only)
C	A history of evaluation in the Australian Public Service	(online only)
D	International approaches to evaluation	(online only)
	References	365

Abbreviations

ABS	Australian Bureau of Statistics
ACF	Administration for Children and Families (United States)
AES	Australian Evaluation Society
AIATSIS	Australian Institute of Aboriginal and Torres Strait Islander Studies
AIFS	Australian Institute of Family Studies
AIHW	Australian Institute of Health and Welfare
ANAO	Australian National Audit Office
ANZSOG	Australia and New Zealand School of Government
APS	Australian Public Service
ATSIC	Aboriginal and Torres Strait Islander Commission
BLADE	Business Longitudinal Analysis Data Environment
CEE	Centre for Evaluation Excellence
CIRCA	Cultural and Indigenous Research Centre Australia
COAG	Council of Australian Governments
CSRE	Centre of Social Research and Evaluation (New Zealand)
Cwlth	Commonwealth
DEG	Dharriwaa Elders Group
DET	Department of Education and Training
DEX	Data Exchange
DFAT	Department of Foreign Affairs and Trade
DHHS	Department of Health and Human Services
DIIS	Department of Industry, Innovation and Science
DIPA	Data Integration Partnership for Australia
DoF	Department of Finance
DOI	Department of the Interior (United States)
DPMC	Department of the Prime Minister and Cabinet
ECPF	Enhanced Commonwealth Performance Framework

FMIP	Financial Management Improvement Program
FNIHCC	First Nations and Inuit Home and Community Care Program
GHS	Get Healthy Information and Coaching Service
HIPPY	Home Interaction Program for Parents and Youngsters
HHS	Department of Health and Human Services (United States)
HM Treasury	Her Majesty's Treasury (United Kingdom)
HREC	Human Research Ethics Committee
IAS	Indigenous Advancement Strategy
IETA	Indigenous Evaluation Threshold Assessment
IUIH	Institute for Urban Indigenous Health
JCB	Job Commitment Bonus
KME	Kaupapa Māori evaluation
KPIs	Key Performance Indicators
LSIC	Longitudinal Study of Indigenous Children
MADIP	Multi-Agency Data Integration Partnership
MoH	Ministry of Health (New South Wales)
NACCHO	National Aboriginal Community Controlled Health Organisation
NATSILMH	National Aboriginal and Torres Strait Islander Leadership in Mental Health
NCARA	NSW Coalition of Aboriginal Regional Alliances
NDIS	National Disability Insurance Scheme
NGOs	non-government organisations
NHMRC	National Health and Medical Research Council
NIAA	National Indigenous Australians Agency
NPP/s	New Policy Proposal/s
OCHRE	Opportunity, Choice, Healing, Responsibility, Empowerment
ODE	Office of Development Effectiveness
OEA	Office of Evaluation and Audit
OECD	Organisation for Economic Cooperation and Development
OIPE	Office of Indigenous Policy Evaluation
ONDC	Office of the National Data Commissioner
PBS	Portfolio Budget Statements

PGPA	<i>Public Governance, Performance and Accountability (Act)</i>
PHN	Primary Health Network
PMCSA	Prime Minister’s Chief Scientific Advisor (New Zealand)
RCTs	randomised controlled trials
RFQ	Request for Quote
RFT	Request for Tender
RSAS	Remote School Attendance Strategy
SPRC	Social Policy Research Centre
TECs	Tribal Epidemiology Centers
UK	United Kingdom
UNDRIP	United Nations Declaration on the Rights of Indigenous Peoples
US	United States

OVERVIEW

Key points

- After decades of developing policies and programs designed to improve the lives of Aboriginal and Torres Strait Islander people, we still know very little about their impact, or how outcomes can be improved.
- Evaluation can answer questions about policy effectiveness, but both the quality and usefulness of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are lacking.
 - Evaluation is often an afterthought rather than built into policy design (and this can affect data collection and evaluation design and result in evaluations that tell you very little).
 - Many evaluations focus on the wrong things (compliance rather than measuring impact, which means findings are often not useful).
 - Aboriginal and Torres Strait Islander people have minimal input into evaluations.
- There is also no whole-of-government approach to evaluation priority setting. And while policy makers agree that evidence is critical for good policies, in practice there is little reliance on evaluation evidence when designing or modifying policies.
- The draft Indigenous Evaluation Strategy (the Strategy) sets out a new approach. It provides a whole-of-government framework for Australian Government agencies for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people.
- The Strategy puts Aboriginal and Torres Strait Islander people at its centre. To achieve better outcomes, what Aboriginal and Torres Strait Islander people value, their expertise, and lived experiences needs to be reflected in what is evaluated, how evaluation is undertaken, and the outcomes policies and programs seek to achieve.
- The Strategy sets out evaluation principles and provides principles-based guidance for Australian Government agencies for selecting, planning, conducting and reporting evaluations. The principles-based guidance aims to 'lift the bar' on the quality of evaluations and improve their usefulness.
- An Office of Indigenous Policy Evaluation (OIPE) is proposed to monitor agencies' performance against the Strategy, provide evaluation leadership, and identify potential cross-agency/topic evaluations. A central clearinghouse for the body of evidence on the effectiveness of policies and programs affecting Aboriginal and Torres Strait Islander people would improve the accessibility of evaluation evidence.
- The Strategy also sets out a whole-of-government approach to evaluation priority setting, including the establishment of an Indigenous Evaluation Council to partner with the OIPE to identify evaluation priorities.
- The case for central evaluation leadership and oversight is wider than the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. Longer term, a new independent agency — a Centre for Evaluation Excellence — could be established to provide evaluation leadership and external oversight for all social and health policy evaluations across Australian Government agencies. If such a Centre was established, the OIPE could move to the Centre as a standalone branch (with the Indigenous Evaluation Council continuing its role).

Overview

The role of governments is to put policies in place to improve outcomes for people and the community. But there is always a degree of uncertainty about how policies will play out in practice. So while well-intended, policies often ‘rest on hypotheses rather than infallible laws’ (Marcel 2015). And policies and programs that do not work well let people down. This is a deficit, compounded by past policy failures, that strikes hard at policies and programs affecting Aboriginal and Torres Strait Islander people. It is reflected in the persistent and yawning gap in outcomes between Aboriginal and Torres Strait Islander people and non-Indigenous Australians.

For decades there have been calls to better understand how policies and programs are affecting Aboriginal and Torres Strait Islander people. But while governments have designed and modified hundreds of policies and programs for Aboriginal and Torres Strait Islander people, and there has been extensive reporting on outcomes for Aboriginal and Torres Strait Islander people against the Closing the Gap targets, the evidence about what works and why, remains thin. And yet it is evidence that is the key to designing policies that will improve the lives of Aboriginal and Torres Strait Islander people.

Empowered Communities (and many others) told the Commission that what was needed was greater insight on policy and program effectiveness:

We need to know what is working and if policy and programs which apply to Indigenous individuals, families and communities are having an impact. We also want greater insight into why policy or program implementation is not effective, and we need early opportunities for correction or reinvestment of funds and effort to ensure that funding is directed to where it is needed most. (sub. 41, p. 5)

Evaluation — the focus of this report — is about answering those questions. Evaluation is the systematic assessment of a policy or program’s design, implementation and outcomes. It is about understanding what governments are doing right, what they are getting wrong and where they can do better. The need for rigorous evaluation of social and health policy has been described as ‘the quiet movement to make government fail less often’ (Leonhardt 2014).

But for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people to make government policies and programs fail less often, and achieve the goal of improving the lives of Aboriginal and Torres Strait Islander people, they need to be high quality and useful. And to be high quality and useful, evaluations need to draw on Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. This is a fundamental shift from the way evaluations of policies and programs affecting

Aboriginal and Torres Strait Islander people are currently undertaken. As one participant to this project said:

... [Aboriginal and Torres Strait Islander] peoples (First Nations Peoples) often know what works and why or why not and what the solutions and proactive steps will be for the community, but they are often not asked, listened to or this gets lost in translation and implementation if they are heard at all. (Liz Curran, sub. 2, p. 1)

To achieve better policy outcomes, Aboriginal and Torres Strait Islander people need to be at the centre. What they value, their expertise and lived experiences 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 those policies have limited value and little prospect of success.

The approach to evaluating policies and programs affecting Aboriginal and Torres Strait Islander people needs to change.

1 What we have been asked to do and our approach

The Commission has been asked to develop a whole-of-government (Australian Government) evaluation strategy for policies and programs affecting Aboriginal and Torres Strait Islander people.

Improving outcomes for Indigenous Australians depends on agencies with responsibility for policies and programs affecting Indigenous Australians undertaking meaningful evaluations. The Commission is to develop a strategy to guide that evaluation effort. (Frydenberg 2019)

The evaluation strategy is to include:

- a principles-based framework for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people
- priorities for evaluation
- an approach for reviewing agencies' conduct of evaluations against the strategy.

How we approached the task

The objective of any policy — in this case a whole-of-government Indigenous Evaluation Strategy (the Strategy) — is to improve the wellbeing of the community overall. Because all policies have costs, to be convinced that a policy will 'make a difference', it is necessary to demonstrate that the benefits of the policy outweigh the costs to the Australian community. The Commission's approach to developing the Strategy involved looking at the strengths and weaknesses of the current arrangements and how a Strategy could improve evaluation, and benefit the lives of Aboriginal and Torres Strait Islander people, and Australians more generally, in light of the costs.

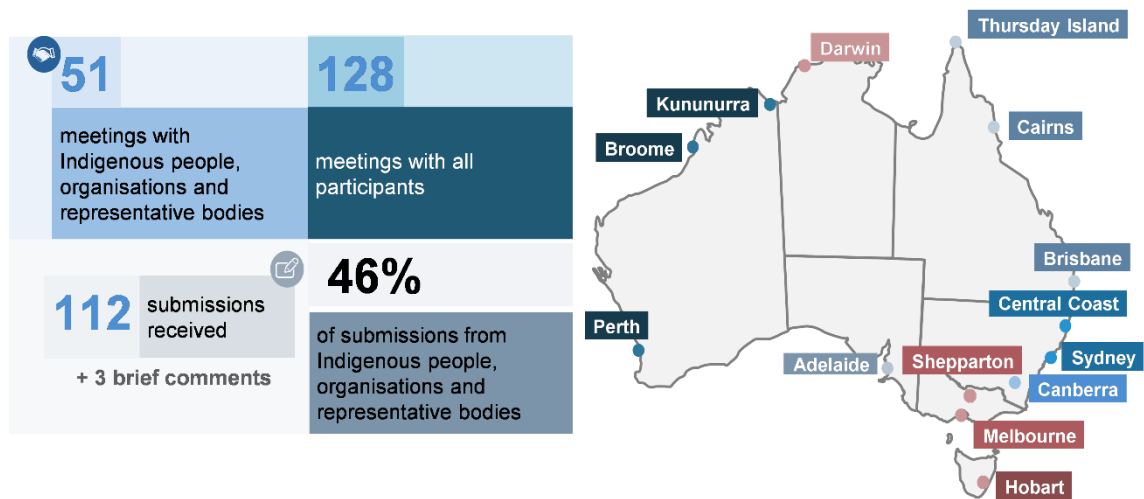
The Strategy covers Indigenous-specific policies and programs as well as mainstream policies and programs that affect Aboriginal and Torres Strait Islander people (the majority of Australian government spending is on mainstream policies and programs).

The Strategy is informed by the literature on good practice evaluation approaches and methods, other evaluation strategies, ethics and research guidelines and standards and strategies, and governance arrangements for evaluation adopted in other similar countries (New Zealand, Canada, the United States and the United Kingdom).

We consulted widely on the Strategy, including with Aboriginal and Torres Strait Islander people, communities and organisations, Australian Government agencies, state and territory and local governments and individuals responsible for administering, delivering and evaluating policies and programs (figure 1).

We held two roundtables — one on ‘Objectives, Principles and Defining Success’ and another on ‘Culture, Capability and Governance’ — to test our thinking on aspects of the Strategy with Aboriginal and Torres Strait Islander organisations, academics and government agencies. We also sent a questionnaire to all Australian Government agencies seeking information on evaluation policies and practices and details of evaluations undertaken by agencies in the previous three years.

Figure 1 **The Commission engaged widely across Australia**



2 The current state of evaluation

Evaluation is currently decentralised within the Australian Public Service (APS). Over time evaluation arrangements have been formalised and centrally supported, then devolved to agencies.

There are no formal requirements for evaluation under the *Public Governance, Performance and Accountability Act 2013* (Cwlth) (PGPA Act). There is currently no central unit providing leadership on or overseeing evaluation, and there is no Australian Government-wide evaluation strategy or policy to guide evaluation across the APS. There is also no co-ordination of evaluation at the whole-of-government level, including for setting evaluation priorities (although some agencies have their own criteria for what to evaluate).

What this means is that the extent to which evaluations are undertaken, and the support agencies provide for evaluation, varies across the APS. Some agencies have their own evaluation strategies and policies, and dedicated evaluation units to manage and oversee evaluation (examples include the National Indigenous Australians Agency and the Department of Foreign Affairs and Trade). Other agencies devolve responsibility for prioritising and managing evaluation to the areas responsible for policies and programs.

There is also no requirement for Australian Government agencies to publish evaluations, nor is there a centralised register or repository for evaluations.

The recent Independent Review of the APS, noting the piecemeal approach to evaluation across the APS, said that this approach ‘diminishes accountability and is a significant barrier to evidence-based policy-making’ (DPMC 2019e, p. 221). The Review, with the aim of embedding a culture of evaluation and learning, recommended a central evaluation function be established. The Australian Government has agreed to establish a small team in the Department of Finance to help build evaluation expertise and practice and develop guidelines to ensure systematic evaluation within the existing performance frameworks. It has also agreed to establish an evaluation profession (DPMC 2019d, p. 22). The implementation of the reforms, however, has been put on hold as the APS responds to the COVID-19 pandemic.

Agencies agree evidence should inform policy, but practice is different

While there is strong agreement across Australian Government agencies that policy should be informed by evidence about what works, we heard that in practice this is often not the case. One government agency told the Commission that a lot of program evaluations provide headline metrics, but not relevant findings and that ‘evaluation is often seen as a cost, not a benefit’. Many said that evaluations are often treated as compliance exercises and the focus is often on measuring outputs rather than impacts, outcomes or effectiveness.

We also heard that:

- evaluation findings are often not timely enough to provide answers to policy questions
- some evaluations are of little value because they have not asked the right evaluation questions
- funding and timeframes for evaluations can be inadequate and this compromises the quality and usefulness of evaluation findings.

That said, some government agencies did point to examples where evaluation evidence had informed policy decisions. One example is the School Enrolment and Attendance Measure, which provided social work support to ensure children enrolled in and regularly attended school (Goldstein and Hiscox 2018, pp. 1–2). This program was stopped after several evaluations, including a randomised controlled trial, found no evidence that it was achieving its main objective of increasing school attendance (DPMC 2018a, p. 90).

Concerns about evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people ...

Participants' concerns about the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people were in four main areas:

- the lack of evaluation (including that few evaluations of mainstream policies and programs examine impacts on Aboriginal and Torres Strait Islander people) and the ad hoc way that evaluations are undertaken
- the quality of evaluations, including that the methodologies adopted for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are often not rigorous and evaluation is not integrated into the policy cycle (rather it is often an afterthought)
- the absence of engagement with Aboriginal and Torres Strait Islander people in determining what is evaluated, and in the design, conduct and interpretation of evaluations
- the usefulness of evaluations, including that evaluations are often not effective in bringing about change, they focus more on accountability than on finding ways to improve outcomes for Aboriginal and Torres Strait Islander people, and they tend to focus on policy or program specific questions rather than cross-agency or cross-topic lessons (box 1).

Box 1 **What participants said about the current state of evaluation**

Cape York Institute:

... there is a lack of evaluation, where evaluations are conducted they are often not very useful in terms of driving change, and evaluations of interventions that show clear and convincing outcomes in Indigenous affairs are few and far between. ... evidence about 'what works', including for whom, under what circumstances, at what cost and why, remains scant. ... Given that the total spend on Indigenous Affairs in Australia is nearing \$35 billion annually it seems perverse and a measure of the problems, that under the current 'system' monitoring and evaluation does not improve results. (sub. 69, pp. 3, 12)

Australian Institute of Aboriginal and Torres Strait Islander Studies:

Evaluation of policies and programs and local outcomes have been undertaken in an ad hoc way, constrained by the parameters of a particular program or activity, such that there is no overarching logic that provides a picture of the impact that government is having in Aboriginal and Torres Strait Islander peoples' lives. (sub. 72, p. 6)

Maggie Walter:

... evaluations only superficially reflect Aboriginal and Torres Strait Islander people's perspectives because input is only sought at the participant level. By that time, important decisions have already been made about the evaluation process — what to evaluate, how to evaluate, the evaluation logic, key outcomes and objectives. Indigenous voice is denied and the current policy mindset is perpetuated. We need Indigenous voice up-front and helping to make decisions from the outset. (sub. 112, p. 2)

AbSec — NSW Child, Family and Community Peak Aboriginal Corporation:

In our experience, government agencies focus on external evidence and impose external solutions on Aboriginal communities, without sufficient regard for their appropriateness and limited evaluation of their effectiveness for local Aboriginal communities. Where evaluations have occurred, they have focused on the priorities and perspectives of government agencies, rather than those of Aboriginal communities and service users. In short, the intended beneficiaries of Aboriginal policy and programs tend to be marginalised in the design, delivery and evaluation of those programs. (sub. 9, p. 4)

Queensland Indigenous Family Violence Legal Service:

... the current evaluation processes are 'top down', have had no input into the actual evaluation framework from ACCO's [Aboriginal Community Controlled Organisations] or Aboriginal and Torres Strait Islander people and are driven by the nature of party politics. (sub. 25, p. 4)

National Aboriginal Community Controlled Health Organisation:

There is little evidence that evaluation is being used in the Indigenous Affairs policy cycle. To the extent it occurs, it appears to be done on an ad hoc basis without any strategy, and much of it seems to be ignored. Evaluations are dominated by the private sector who do not have any real relationship with Aboriginal and Torres Strait Islander peoples. (sub. 95, p. 1)

Tony Dreise et al.:

Too often, evaluations of policies affecting First Nations peoples have been poorly undertaken. ... Too often evaluations have been designed in such a way that causal inference is impossible. (sub. 33, p. 10)

... align with some of our questionnaire results

The results from our questionnaire to Australian Government agencies (46 agencies responded, box 2) align with what participants told us about evaluations affecting Aboriginal and Torres Strait Islander people. We received details on just over 100 evaluations undertaken since 2016-17 that either mentioned or provided results for Aboriginal and Torres Strait Islander people. The results (based on this sample of evaluations) show that:

- Aboriginal and Torres Strait Islander people have minimal input into evaluations
 - one-fifth of evaluations of Indigenous-specific policies and programs reported engaging with Aboriginal and Torres Strait Islander people in evaluation decision making
 - one out of 47 evaluations of mainstream policies and programs involved Aboriginal and Torres Strait Islander people in evaluation planning and decision making
 - two-thirds of mainstream policy or program evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people did not report engaging with Aboriginal and Torres Strait Islander people
- evaluations relied heavily on qualitative methods, including surveys, interviews, consultation meetings and focus groups. Qualitative methods were the main method used in around one-quarter of evaluations
- most evaluations attempted to measure impact but were often not robust enough to do so, due to limited data and/or the absence of a control group (which suggests that evaluation was not planned for early or integrated into the policy cycle). Very few evaluations included cost-benefit analysis
- two-thirds of the evaluations of mainstream policies and programs did not consider the impacts of policies and programs on Aboriginal and Torres Strait Islander people
- most evaluations were policy or program specific — just three Indigenous-specific policy evaluations and one mainstream program evaluation covered multiple agencies or service areas.

Our review of the evaluation evidence base for policies and programs affecting Aboriginal and Torres Strait Islander people (noting that many evaluations are not published) also found that despite decades of developing new policies and programs and modifying existing programs, there are significant gaps in the evidence. We still know very little about the impact of policies and programs in many policy areas, or how outcomes could be improved. Many questions remain unanswered.

Box 2 **Some results from our information request**

In late 2019, the Commission sent an evaluation questionnaire to 182 Australian Government agencies. The questionnaire asked agencies about their evaluation policies/strategies and evaluation practice. Agencies were also asked to provide copies of (or links to) evaluations published over the period 2016-17 to 2019-20 and details of any unpublished evaluations conducted over the same period.

Forty-six agencies responded to the questionnaire — 13 departments and 33 agencies (14 medium to large and 17 small mainstream agencies, and two Indigenous-specific agencies). Agencies provided information on 207 evaluations conducted between 2016-17 and 2019-20. Two-thirds of the evaluations were published. Agencies also mentioned (but did not provide details on) a further 100 unpublished evaluations.

Of the 207 evaluations provided by agencies, 102 mentioned or provided results for Aboriginal and Torres Strait Islander people (47 evaluations were for mainstream policies or programs and 55 for Indigenous-specific policies or programs).

Evaluation effort (measured by the number of evaluations undertaken) varies considerably across agencies. And effort does not seem to be correlated with agency size. Publication practices for evaluation reports also vary across agencies — ten per cent of agencies said they ‘always’ published evaluation reports and around a third ‘very often’.

Some of the insights from the evaluations of policies and programs mentioning or providing results for Aboriginal and Torres Strait Islander people were that:

- about a quarter of the evaluations provided related to public and community health services, 15 per cent to school education, 14 per cent to labour and employment and 12 per cent to community and environment services. Areas of significant government expenditure where there were few or no evaluations provided were social security payments and healthcare subsidies and support (including Medicare and the Pharmaceutical Benefits Scheme)
- about two-thirds of the evaluations used mixed methods (quantitative and qualitative data)
- fewer than 10 per cent of the evaluation reports mentioning or providing results for Aboriginal and Torres Strait Islander people included information on how the evaluation findings or recommendations were used by the commissioning agency
- about 60 per cent of evaluations of both mainstream and Indigenous-specific policies or programs that mentioned or provided results for Aboriginal and Torres Strait Islander people were undertaken by external consultants
- ethics assessment was undertaken for a small proportion of evaluations — around one in five mainstream program evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people and one in four Indigenous-specific program evaluations.

Source: Productivity Commission analysis based on an evaluation questionnaire (appendix B).

3 A new approach — an Indigenous Evaluation Strategy

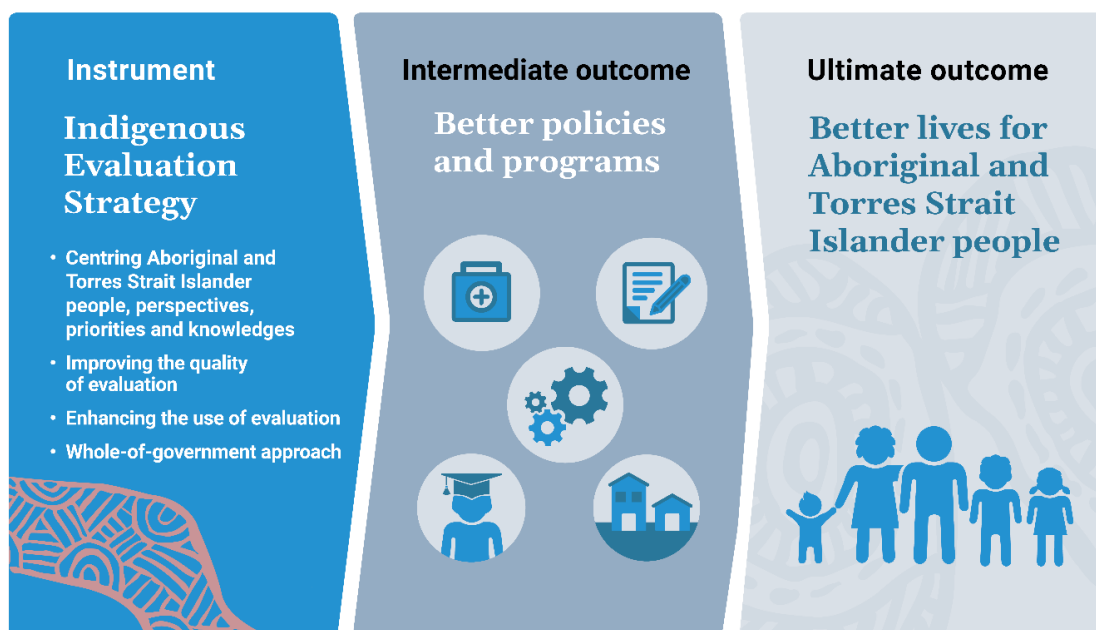
There was widespread agreement that the objective of the Indigenous Evaluation Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people.

For the Strategy to achieve this objective, it needs to be more than a document that sets out objectives, principles and priorities for selecting and conducting evaluations. 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.

We identified four ways an Indigenous Evaluation Strategy, via better policies and programs, could improve the lives of Aboriginal and Torres Strait Islander people (figure 2). They are by:

- centring Aboriginal and Torres Strait Islander people, perspectives, priorities, and knowledges in evaluations of policies and programs that affect them
- ‘lifting the bar’ on the quality of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people
- enhancing the use of evaluations to inform policy and program design and implementation, including by supporting a culture of evaluation and building a body of evidence and data on the effectiveness of policies and programs
- promoting a whole-of-government approach to priority setting and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.

Figure 2 **The objective of the Strategy is to improve the lives of Aboriginal and Torres Strait Islander people**



Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges

Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges into program and policy evaluation planning, design, implementation and priority setting is critical for improving outcomes for Aboriginal and Torres Strait Islander people.

Many participants called for Aboriginal and Torres Strait Islander people’s engagement in evaluation design and implementation, and for the Articles in the *United Nations Declaration on the Rights of Indigenous Peoples*, including the right to self-determination, to be recognised in the Strategy (box 3). While self-determination means different things to different people, the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) explained the principle of self-determination as requiring ‘... Indigenous peoples be involved in decisions that affect them, including the design, delivery and evaluation of government policies and programs’ (sub. 72, p. 5).

As noted earlier, many of the concerns about the quality of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people were in the context of the lack of engagement with Aboriginal and Torres Strait Islander people — ‘somebody else tells our story for us’, and the lack of regard for Indigenous knowledges.

Effectively engaging with Aboriginal and Torres Strait Islander people and communities in policy design and evaluation will ensure that what they value, their perspectives, expertise and lived experiences are addressed and reflected in outcomes. As the National Mental Health Commission put it, a principle of incorporating engagement with people directly affected by a policy, program or service encourages:

... an approach to evaluation that moves beyond a simple focus on inputs and outputs, to a focus on experiences and outcomes. This approach would provide a deeper and more nuanced understanding of the impacts of policies, programs and services on the communities that are affected by them, and of the factors that drive positive outcomes. (sub. 19, p. 3)

Box 3 Indigenous people, perspectives, priorities and knowledges are critical for the Strategy's success

Institute for Urban Indigenous Health:

This submission advocates for a recalibration of focus in program design, delivery and evaluation which ... builds Indigenous perspectives and leadership into the design and delivery of all parts of the policy/program cycle and at all levels (national, state and local/regional) as a necessary precursor to improving outcomes for Aboriginal and Torres Strait Islander peoples. (sub. 53, p. 3)

National Health Leadership Forum:

The NHLF supports an [Indigenous Evaluation Strategy] that is centred on a principles-based evaluation framework ... This framework should be underpinned by the United Nations Declaration on the Rights of Indigenous Peoples, which mandates the involvement of Indigenous Peoples as equal partners in priority setting and decision-making regarding policies and programs that affect them. (sub. 59, p. 3)

Queensland Indigenous Family Violence Legal Service:

The primary objectives to be achieved, in a strategy aimed at evaluating policies and programs affecting Aboriginal and Torres Strait Islander peoples are: ensuring that Aboriginal and Torres Strait Islander peoples are actively involved in decision making affecting them and play a valuable part in the co-design and all decision making stages; ensuring that Aboriginal and Torres Strait Islander peoples' voices are heard and not used merely as a 'tick and flick' mechanism; to incorporate Aboriginal and Torres Strait Islander knowledge, priorities and perspectives during evaluation. (sub. 25, p. 3)

The Western Australian Government:

... the strategy should aim to maximise Aboriginal engagement in evaluation design and implementation, by focusing on genuine partnership and co-design ... (sub. 74, p. 1)

The Smith Family:

An Indigenous Evaluation Strategy must rest on the principle of self-determination ... This principle implies the importance of Aboriginal and Torres Strait Islander people fully participating in the development, implementation and management of programs and policies that affect them, including the development of an Indigenous Evaluation Strategy and evaluation activities which flow from it. (sub. 56, p. 3)

Catholic Care NT and the Australian Centre for Social Enterprise:

Understanding Indigenous life experiences and knowledges, which vary across communities and nations, should frame the development of outcomes and the consequent evaluation. (sub. 75, p. 6)

A focus on experiences and outcomes involves recognising the diverse cultures (languages, knowledge systems and beliefs) and histories of Aboriginal and Torres Strait Islander people, and the impacts of contemporary and historical policies and programs on the wellbeing of Aboriginal and Torres Strait Islander people and communities.

Genuine partnerships with Aboriginal and Torres Strait Islander people can also strengthen evaluation capability and facilitate a sense of joint ownership. And it can strengthen trust and the cultural capability of non-Indigenous evaluators and those delivering policies or programs (box 4 provides an example).

Evaluations of successful policies and programs reinforce the importance of Aboriginal and Torres Strait Islander people and communities in the design, development and delivery of policies and programs (Osborne, Baum and Brown 2013; QPC 2017). The evidence also shows that what works is engagement through partnerships within a framework of self-determination and strategies that address power inequalities and genuine efforts to share power (Hunt 2013).

We are recommending that the overarching principle of the Indigenous Evaluation Strategy be centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (figure 3). This principle recognises that Australian Government agencies, when evaluating, need to draw on Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges if outcomes are to improve. As figure 4 outlines, what this means in practice is engaging with Aboriginal and Torres Strait Islander people on what to evaluate (evaluation priorities), evaluation questions, approaches, methods and practices, and the interpretation of evaluation findings.

Box 4 **A qualitative evaluation of a Mobile Dialysis Truck**

Rates of end-stage kidney disease among Aboriginal and Torres Strait Islander people in remote areas are disproportionately high. However, because haemodialysis is not offered in many remote areas, Aboriginal and Torres Strait Islander people have to leave their Country (with its traditions and supports) and relocate to metropolitan or regional centres to receive treatment. This disrupts the kinship and cultural ties that are important for their wellbeing.

The South Australian Mobile Dialysis Truck visits remote communities and allows patients living in regional or metropolitan centres to return home for significant community events, spend time on Country with family and friends and have dialysis on Country.

An evaluation of the Mobile Dialysis Truck involved face to face semi-structured interviews with 15 Aboriginal and Torres Strait Islander dialysis patients and 10 nurses who had attended trips across nine dialysis units. The interviews were conducted using a 'yarning' method (box 7) to collect information in a culturally safe way, and in a relaxed, informal setting. Realist evaluation methodology and thematic analysis were used to understand patient and nursing experiences with the Mobile Dialysis Truck.

Aboriginal and Torres Strait Islander patients said the consequences of leaving Country included grief and loss. The identified benefits of the service included: the ability to fulfil cultural commitments, minimisation of medical retrievals from patients missing dialysis to return to remote areas, improved trust and relationships between patients and staff, and improved patient quality of life. The Mobile Dialysis Truck was also found to be a valuable cultural learning opportunity for staff (with this form of health service potentially improving cultural competencies for nursing staff who provide regular care for Aboriginal and Torres Strait Islander patients).

The Mobile Dialysis Truck was found to improve the social and emotional wellbeing of Aboriginal and Torres Strait Islander people who had to relocate for dialysis. It was also found to build positive relationships and trust between metropolitan nurses and remote patients.

Some of the identified barriers to attending the Mobile Dialysis Truck on Country included infrequent trips, ineffective trip advertisement, poor patient health and a lack of appropriate or unavailable accommodation for staff and patients.

The evaluators identified a number of potential areas for future research, including collecting:

- data from patients who had not been on trips, to elicit their experiences and compare data on quality of life
- quantitative data on hospital admissions and Royal Flying Doctor Service retrievals (to see if they were reduced in the presence of the dialysis truck)
- biomedical data on dialysis encounters (to see if adherence to dialysis and associated management regimes improved for patients who have spent time on the dialysis truck).

Source: Conway et al. (2018, pp. 1, 11).

Figure 3 **Guiding principles for the Indigenous Evaluation Strategy**

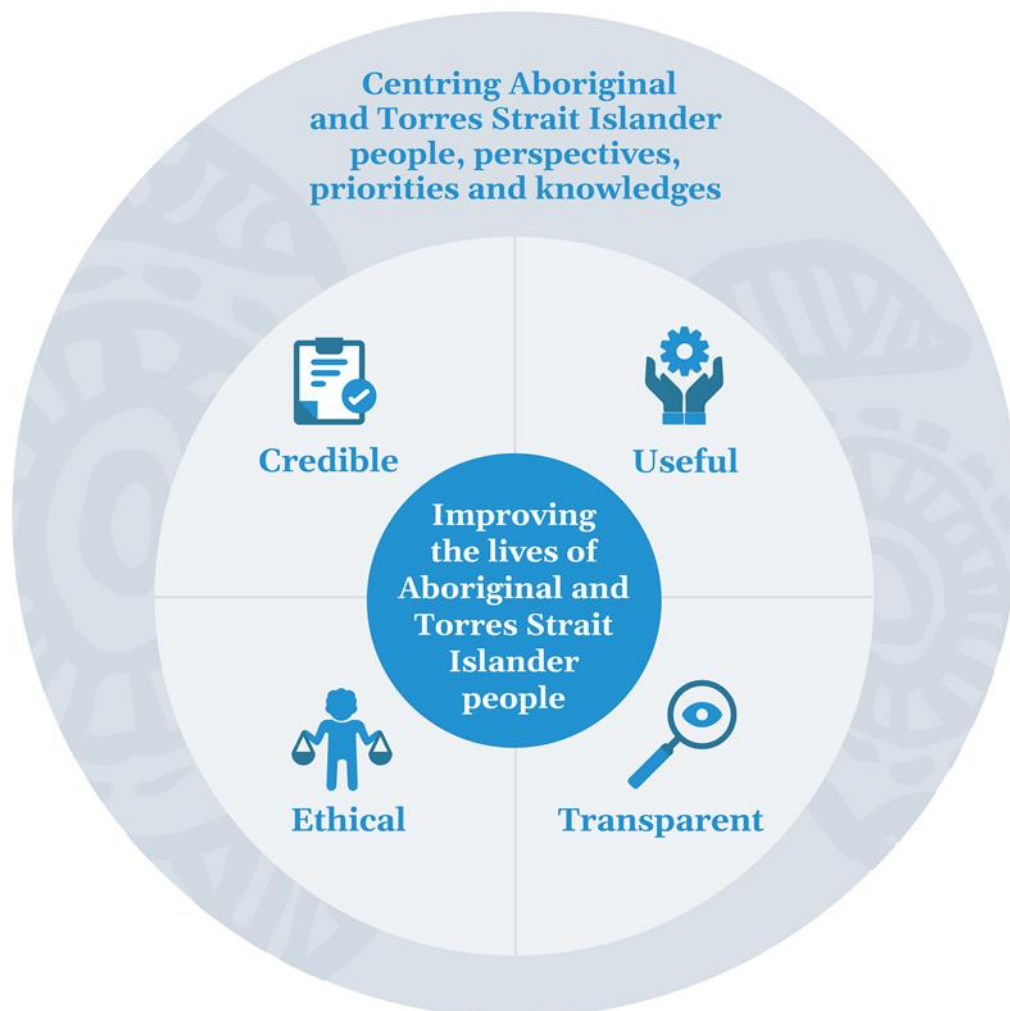


Figure 4 Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in practice

What to evaluate	Aboriginal and Torres Strait Islander people are engaged to decide what policies and programs achieve the greatest impact on their lives and should be subject to rigorous evaluation.
Evaluation planning, design and conduct	- Aboriginal and Torres Strait Islander people have the option of being partners in all stages of evaluation. This includes engaging with Aboriginal and Torres Strait Islander people on evaluation questions, evaluation approaches and methods, and the interpretation of evaluation findings. - Sufficient time and resources are allowed for meaningful engagement with Aboriginal and Torres Strait Islander people during evaluation. - Engagement between commissioners, evaluators, participants and users is respectful of differences, mutually beneficial and undertaken in culturally safe ways. - Evaluators have the necessary skills and experience working in partnership with Aboriginal and Torres Strait Islander people to undertake the evaluation. - Evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are responsive to the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people and communities. - Mainstream policies and programs routinely consider impacts on Aboriginal and Torres Strait Islander people — and evaluate those where the impact is considered significant. - Evaluation design and reporting reflect the diversity of Aboriginal and Torres Strait Islander people, their priorities and experiences.
Reporting evaluation findings	- Aboriginal and Torres Strait Islander people are engaged in translating evaluation findings to ensure they are meaningful, accessible and useful for communities and decision makers. - Evaluation reports describe how evaluators engaged with Aboriginal and Torres Strait Islander people during the evaluation process. - Evaluators and commissioners ensure that evaluation findings are communicated back to the Aboriginal and Torres Strait Islander people, communities and organisations that participated.
Building capability and a culture of evaluation	- 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. - Evaluation processes seek to build capability among Aboriginal and Torres Strait Islander evaluators, organisations and communities. - Agencies provide opportunities or encourage staff to pursue opportunities to strengthen their cultural capability. - Agencies are open to review and feedback from Aboriginal and Torres Strait Islander people and organisations about their evaluation practices.

Centring is also about valuing Aboriginal and Torres Strait Islander knowledges, cultural beliefs and practices, and creating a sense of ownership and building capability among Aboriginal and Torres Strait Islander evaluators, organisations and communities. And it is about non-Indigenous evaluators having the necessary skills and experience to work in

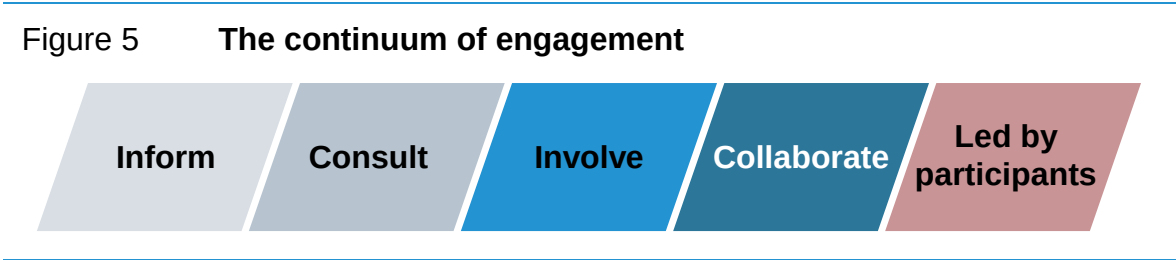
partnership with, and to draw on the knowledges of, Aboriginal and Torres Strait Islander people.

For mainstream services, centring is about agencies considering the potentially different impacts of policies and programs on Aboriginal and Torres Strait Islander people when designing policies and programs and undertaking evaluations. This means: planning early so the right data (including baseline data) are collected to track and report on outcomes for Aboriginal and Torres Strait Islander people; ensuring the program logic reflects on the potential impacts on Aboriginal and Torres Strait Islander people (and potentially other minority groups); and appropriate evaluation questions are built into the evaluation.

Co-design

Co-design is an important approach to incorporating Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges into evaluations. The term co-design, however, is used and understood to reflect different levels of engagement across the continuum of engagement, although it is more than consultation and usually falls within the collaborate and participant-led end of the continuum (figure 5).

In the context of the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, it is about engaging Aboriginal and Torres Strait Islander people and communities in the design and conduct of an evaluation in a way that is meaningful and engenders respect, empowerment and ownership (Dreise and Mazurski 2018, p. 8).



The Department of Health, reflecting on the four year independent evaluation of the Australian Government’s investment in Aboriginal and Torres Strait Islander Primary Health Care under the Indigenous Australians’ Health Programme,¹ which is co-designed with experts in Aboriginal and Torres Strait Islander health and with local Aboriginal and Torres Strait Islander communities, said:

Co-design provides a tool for system learning and adaptation and for Aboriginal and Torres Strait Islander Australians to have a real voice. It can be used to disrupt the status quo and help people to problem solve collectively and bring about positive change to improve outcomes for

¹ The \$18.6 million evaluation is the largest investment in a single Indigenous health evaluation to date. It is being undertaken by a collaboration that includes Allen + Clarke, Monaghan Dreaming and the Centre for Indigenous Equity Research (Central Queensland University) (Australian Government Department of Health, sub. 35, Attachment A, pp. 1-2).

Aboriginal and Torres Strait Islander Australians. To date the evaluation team, the Department, and the Health Sector Co-design Group have reflected that co-design:

- Is a process
- Takes time
- Is about relationships
- Requires trust and trustworthiness. (sub. 35, Attachment A, pp. 7–8)

Co-design can improve transparency — participants are aware of (and engaged in deciding on) evaluation questions, evaluation approaches and methods and how to disseminate evaluation results. And because co-design is collaborative in nature, as Bradwell and Marr said, ‘[it] shifts power to the process ...’ (2008, p. 17). Co-design also recognises that evaluation is closely connected to the design, purpose and implementation of a policy or program itself.

However, co-design, can be resource and time intensive, and those involved need to have access to the information, skills and capacities to be able to participate effectively. A reflection from the NSW Government’s OCHRE evaluation was that the co-design process was complex for evaluators and participants alike, and that it was hard to gauge ahead of time the resources and time required for co-design (Dreise and Mazurski 2018, p. 23).

Centring and priority setting

Evaluations should also be undertaken in the areas that are valued most highly by Aboriginal and Torres Strait Islander people. As part of the proposed governance arrangements for the Strategy, we are recommending an Indigenous Evaluation Council, with a majority of Aboriginal and Torres Strait Islander members, to partner with an Office for Indigenous Policy Evaluation (OIPE) to identify whole-of-Australian Government evaluation priorities (further details are provided below). This is in line with the Australian Government’s policy of working in genuine formal partnership with Aboriginal and Torres Strait Islander people and the overarching principle of the Strategy.

Lifting the bar on the quality of evaluations

Quality of evaluations matters

Because evaluations inform decisions and actions that affect the lives of Aboriginal and Torres Strait Islander people and other Australians, they need to be of high quality and produce credible, useful and timely results that can be relied on by decision makers. One of the findings from evaluations — unsurprisingly — is that policies and programs implemented with an evidence base behind them are more likely to have successful outcomes (Bowes and Grace 2014).

Done well, an evaluation will give insights on how a policy or program is being (or has been) implemented and what effect it has, for whom and why. It will also identify how the policy or program can be improved. Done badly, evaluations can produce misleading results and there is a risk that the evidence produced will result in incorrect decisions about policies and programs (such as allowing ineffective or harmful policies to continue). The quality of evaluation matters.

The Strategy sets out core evaluation principles and principles-based guidance for Australian Government agencies for selecting, planning, conducting and using evaluations. The four principles for framing what should occur (reflecting good practice) and how quality of an evaluation will be assessed are — credible, useful, ethical and transparent (with centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges woven through all the principles, figure 3).

Planning early and building evaluation into policy design is the key to quality

Planning early and building evaluation into the design of policies and programs² at their inception is key to quality evaluations. It is important for ensuring evaluations identify the objectives and outcomes of the policy or program and clarify what the evaluation questions should be. Developing a logic model or theory of change (as a way to be clear about the objectives of a policy or program) can result in both better policy planning and evaluation, and ensure that the things that really matter are measured during the evaluation (table 1 sets out questions that should be asked at the planning stage of an evaluation).

Good data are critical for quality evaluations. Building evaluation into the design of policies and programs helps ensure that the data and evidence needed to answer the evaluation questions and to produce credible results are collected (including, for example, the collection of baseline data and finding an appropriate comparison group).

² Before a policy or program is in place, previous evaluations should be reviewed and the evaluation evidence used to shape the design and the way it is implemented.

Table 1 Questions to consider at the planning stage of an evaluation

<i>The main steps</i>	<i>Questions to consider</i>
Define the policy objectives and outcomes to identify what to evaluate	What is the program logic or theory of change? Does it clearly state how the program outcomes are expected to be achieved and how the outcomes are linked to higher level outcomes? Do the intended outcomes align with the priority needs of Aboriginal and Torres Strait Islander people and other stakeholders?
Identify and engage with those who will be affected by the policy or program	Who will be affected by the policy or program? (This should include those involved in delivering the initiative, users/recipients, others affected by the policy or program, and those interested in the evaluation results.)
Identify the evaluation questions	What do policy makers and those affected by the policy or program need to know? (Evaluation questions are best decided on with the main evaluation information users.)
Select an evaluation approach and method(s)	What approaches and methods will answer the evaluation questions? Is the evaluation a formative, process, or impact evaluation, or some combination? Is an economic evaluation required? Are approaches/methods identified for each evaluation question? Are contextual factors that are likely to influence outcomes identified?
Identify data requirements (quality evaluations require good data)	What data are required? What is already being collected/available? If undertaking an impact evaluation, how will baselines be collected and when will impact be measured? What processes need to be put in place to allow the data to be collected? Has consideration been given to the design of data collection methods that are responsive to the needs, rights and cultural safety of respondents?
Identify ethics requirements	What ethics processes are required? Has time been allowed for ethics processes?
Identify resource and governance needs	What level of resourcing is required for the evaluation? Does the evaluation plan allow sufficient time and resources to engage meaningfully with relevant parties at each stage of the evaluation? What governance arrangements (a steering group or peer review) will need to be put in place? What quality control processes will there be?
Conduct the evaluation	Will the evaluation be conducted internally or be externally commissioned? Who will be responsible for tendering, project management and quality assurance? When will the evaluation start and end?
Use and disseminate the evaluation findings	What will the findings be used for? What decisions will they feed into? How will the findings be shared and disseminated? Will it be necessary to present the findings in different formats for different users? What approaches will be used to ensure information is considered during relevant policy decision making?

Source: Adapted from HM Treasury (2011).

Early planning for evaluation is one way to ensure that 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. Important evaluation questions for mainstream policies and programs are — is it working differently for Aboriginal and Torres Strait Islander people, or differently for groups of Aboriginal and Torres Strait Islander people (or other groups)? Why, or why not, might it be working differently for different groups? Are Aboriginal and Torres Strait Islander people accessing the policy or program at the same rate as other Australians? If not, why not? Are there barriers to Aboriginal and Torres Strait Islander people accessing the policy or program? Is the policy or program being delivered in a culturally appropriate way?

Data collection methods and tools used for the general population are not always effective for collecting data on Aboriginal and Torres Strait Islander people and more tailored approaches may be required. An issue for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people is small sample sizes (many data sources do not have sufficient numbers of Aboriginal and Torres Strait Islander people or fail to consistently collect data on Indigenous status), and small sample sizes can limit statistical power. As our information request results showed, many evaluations were not able to answer evaluation questions because of limited data (or the absence of baseline data).

What is important is asking questions early about what data (including what data captures what matters for Aboriginal and Torres Strait Islander people) and sampling strategies are required to measure impacts for Aboriginal and Torres Strait Islander people. Considering sampling issues early means that appropriate sampling strategies can be chosen and data collected (an example of oversampling, a sampling strategy for small sample sizes, is provided in box 5).

Evaluation quality is also strongly tied to the conduct of evaluation, including how evaluators engage with those affected by the policy or program. Developing an evaluation plan as part of the policy or program design enables Aboriginal and Torres Strait Islander people, policy decision makers and others likely to be affected by the policy or program, to be involved early on in the process and to shape policy objectives, outcomes, and evaluation approaches and methods, data and data collection. As the World Bank said, ‘the absence of stakeholder engagement in the definition of the ‘evaluation problem’ risks reducing the utility of an evaluation or even creating resistance to evaluation findings’ (World Bank 2019, p. 12).

Box 5 **Oversampling — the Adult Admitted Patient Survey**

The Adult Admitted Patient Survey asks for feedback from people who have recently been admitted to a New South Wales public hospital. Each year the survey is sent to between 70 000-85 000 people approximately three months after the end of their hospital stay. Participants are randomly selected to complete the survey in order to provide a representative sample across New South Wales.

For the 2014 survey, Aboriginal and Torres Strait Islander patients were oversampled (sampled at a rate higher than the proportionate representation in the total patient population) to ensure that a sufficient number of responses were collected to allow statistical analysis.

The oversampling strategy involved the selection of all eligible Aboriginal and Torres Strait Islander patients (that is, census design). Although Aboriginal and Torres Strait Islander people made up 2 per cent of the population, they represented 10 per cent of survey respondents. (The response rate among Aboriginal and Torres Strait Islander patients was 21 per cent, which was much lower than the response rate for all patients (44 per cent), and was another reason to oversample.)

Responses were then weighted to ensure that the overall survey response rate reflected a response rate that would be observed if patients were sampled proportional to the actual patient population.

Sources: Bureau of Health Information (2015, 2016).

Planning early also means that thought can be given early on to how to reach and engage with particular communities or groups, how to engage in a culturally safe way, and ensure the time and resources are factored in to the evaluation for such engagement. As the Australian Evaluation Society said:

Developing the theory of change, program theory or program logic with Indigenous people and communities helps to underpin the design and methods which are more culturally appropriate to the communities in which the evaluation is being conducted. (sub. 49, p. 17)

Early planning for evaluation can also help build a culture of evaluation, as evaluation becomes embedded in the policy cycle rather than being something that agencies do as an afterthought.

Evaluation types, approaches and methods

There are many different types, approaches and methods that can be used to evaluate policies and programs. The challenge is to evaluate in a way that produces credible, robust and useful results.

The Strategy does not endorse particular evaluation types, approaches or methods. The evaluation type and approach used should depend on the questions that need to be answered (table 2), the kind of policy or program being evaluated, the circumstances under which the policy or program is being implemented, the time and resources allocated to the evaluation, and the types of data needed to answer the evaluation questions (including what information

is considered credible by those affected by the policy or program). Evaluations should be based on the best possible approaches and methods given the data, time and resources available.

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>
Formative or process evaluations — these evaluations are undertaken early in the development or implementation of a policy or program.	How is the policy or program delivered? Is the program being delivered as intended? Is the policy or program appropriately targeted? How effective has implementation been so far? What are the policy or program's strengths and weaknesses? 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.
Summative, outcome or impact evaluations — these evaluations judge the overall merit, worth and impact of a policy or program.	What difference did the policy or program make? Has the policy or program achieved its objectives? Does the evidence support the theory? Has the policy improved outcomes? If so, by how much? 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.
Economic evaluations — these evaluations quantify the net benefit of a policy or program.	Do the benefits justify the costs, or was it worth it? 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.

More than one type of evaluation can be used, and in fact, different types of evaluations can complement each other. For example, a formative or process evaluation undertaken in the early years of a policy or program (while working out start-up problems) could be followed by an impact evaluation after the policy or program is established (to judge its merit or worth).

Evaluation rigour can also be enhanced by employing triangulation, and using mixed methods (combining quantitative and qualitative methods). Qualitative data can help answer 'why' questions and provide a more complete picture. As Colleen Hayward said:

Quantitative [data], as you know, is just going to give you the numbers. They don't tell you the story behind it and in my view, it is the stories behind the numbers where the richness actually lies ... (in Smith et al. (2018, p. 36))

The validity of sampling designs and data analysis techniques are critical for technical rigour in impact evaluations. In the context of qualitative methods, rigour refers to findings that represent as closely as possible the experiences of participants (that is, using techniques that ensure you have a representative sample).

Are some approaches more or less suited?

Some participants suggested that approaches to impact evaluation, including experimental approaches such as randomised controlled trials (RCTs), were less suited to understanding policy and program impact for Aboriginal and Torres Strait Islander communities.

RCTs are used widely in areas of medical, health and other scientific research, where treatments are homogenous and defined, and where the nature of experiments are largely under the control of researchers. They are used less in evaluation for social policy where policies and programs are designed to deal with multiple and complex issues and where change is expected to occur in the long term (although their use in social policy is increasing). Figure 6 outlines the conditions under which experimental and quasi-experimental approaches to impact evaluation are more (or less) feasible.

Many of the less suitable conditions for experimental or quasi-experimental approaches (figure 6) will often apply to evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people and communities (including, for example, small sample sizes, the policy or program being difficult to disentangle from other programs, a complex relationship between the expected outcome and the policy or program, or where it may not be practical or ethical to create different comparison and target groups). RCTs can also be expensive which means they need to be restricted to high value policies and programs.

Also, because RCTs can hide asymmetric effects across a study population, especially in heterogeneous samples, evaluators need to carefully consider how and when to use RCTs to evaluate policies and programs affecting Aboriginal and Torres Strait Islander people. That said, there will be circumstances where RCTs are suitable for policies and programs affecting Aboriginal and Torres Strait Islander people (for example Goldstein and Hiscox (2018), and Ward et al. (2013)).

Figure 6 **When to use quantitative impact evaluation**

		More suited if...	Less suited if...
Nature of the policy or program	➔	The initiative is new, or a distinctive change is made to an existing one Activities are standardised for all participants The initiative and its objectives are distinctive	Changes to the initiative are small or incremental Activities vary between participants Many overlapping initiatives contribute to similar objectives
Type of delivery	➔	The initiative can be delivered selectively, or phased in, so that participants can be allocated to treatment and control groups	Selective delivery is undesirable or not possible Delivery is universal or launched simultaneously nationwide
Expected impacts	➔	A direct, causal relationship is expected between an initiative's activities and outcomes Effects are immediate or expected in the near term The expected impact of the initiative is large compared to other influences The influence of extraneous factors can be controlled for	Activities and outcomes are only weakly or indirectly linked Effects are gradual, achieved in the longer term The expected impact of the initiative is small compared to other influences Many complex, interrelated factors influence outcomes that cannot be controlled
Data quality	➔	Baseline data are collected prior to roll out Unit-level data are collected at precise time points Variables are easily quantified	Data are collected retrospectively Data are aggregated and collected at imprecise times Variables are difficult to quantify
Population of interest	➔	A large population exists from which to draw samples	Working with small populations, sub-populations or participant groups

Sources: Ames and Wilson (2016); Rogers (2009b); White, Sabarwal and de Hoop (2014). Adapted from Campbell and Harper (2012, p. 7).

Developmental, realist and participatory evaluation were highlighted by participants as being particularly suited to evaluating policies and programs affecting Aboriginal and Torres Strait Islander people on the basis that they:

- are better suited to complex problems and adaptive learning
- can be more effective at incorporating Aboriginal and Torres Strait Islander knowledges, perspectives and world views into evaluations
- can engage Aboriginal and Torres Strait Islander people throughout the evaluation process, including using culturally sensitive methods, such as yarning (storytelling), ganma (knowledge sharing) and dadirri (listening) (box 6)
- can build participants' evaluation capabilities.

Developmental approaches to evaluation have grown out of observations about how innovation occurs in the corporate sector and the idea that answers are not known but can be discovered through a process of trying things (Gamble 2008). They seek to maximise the success of a policy or program through incremental review and improvement as it is rolled out and rely heavily on program participants' responses to shape learning and improvement.

Realist evaluation seeks to understand the circumstances under which some programs are more or less effective with different groups of people. This approach seeks contextual explanations for why change occurs, while seeking out diversity of participants' experiences.

An example of a realist evaluation is provided in box 7. This evaluation found that the Remote School Attendance Strategy, which engages local people to work with schools and families to support children to attend school (and is mostly rolled out in Aboriginal and Torres Strait Islander communities) did not improve school attendance in all families, and acting in isolation it was unable to improve attendance for families experiencing complex life events. However, the evaluation also found that when staff were appropriately trained, knowledgeable about other services, and were seen as 'the right person for the job', the program worked as an effective referral and coordination point for families (DPMC 2018d).

Box 6 **Yarning and dadirri**

Narratives and stories are a fundamental part of sharing Aboriginal and Torres Strait Islander ways of knowing.

Yarning is a qualitative research method that draws on Aboriginal and Torres Strait Islander people's cultural practice of storytelling as a way of conveying information and for relationship building (it can be one-on-one or in groups). It involves in-depth discussions to gain 'thick description' and offers a culturally safe place for Aboriginal and Torres Strait Islander people to freely talk about their experiences. Yarning positions the evaluator as a listener and learner in the data collection process and respects Aboriginal and Torres Strait Islander people as the authority of their knowledges. It involves reciprocal relationships and is a two-way process of learning and knowledge exchange.

The Victorian Aboriginal Community Controlled Health Organisation described the different types of yarning:

... 'social yarning' (which builds trust and engagement), 'research topic yarning' (a relaxed interview that focuses on information gathering), 'collaborative yarning' (sharing ideas or brainstorming about the research topic), and 'therapeutic yarning' (participant disclosure of emotional or traumatic experiences). The diversity of types of yarning allows it to be applied to various elements within the process of evaluation. (sub. 44, p. 15)

Dadirri is an Aboriginal concept and practice of deep contemplation and listening. The word comes from the Ngan'gikurunggurr and Ngen'giwumirri languages (meaning 'listening'). Dadirri has also been described as both an 'indigenous way of life' and 'a research methodology'; and as a methodological approach, indigenous philosophy informs the investigative process, and ensures ethical behaviour and cultural safety in research design. It is an approach that:

- honours the integrity and fidelity of Aboriginal communities
- requires reciprocity in how researchers relate to, and behave with, Aboriginal people
- encourages researchers to reflect on their own assumptions or bias
- uses Aboriginal world views in which learning is a process of listening and witnessing without judgement or prejudice, and of being responsible for self in relationships with others in the listening and learning process
- upholds Aboriginal world views so that the activity of learning introduces a responsibility to act with integrity and fidelity to what has been learnt (Atkinson 2002, p. 20)

Sources: Atkinson (2002, pp. 15, 20); Geia, Hayes and Usher (2013, p. 16); Leeson, Smith and Rynne (2016, p. 8).

Box 7**A realist-informed and participatory approach to better understand school attendance**

The Remote School Attendance Strategy (RSAS) is a place-based initiative focused on supporting school attendance. The Commonwealth and state governments deliver a range of programs under the initiative to support school engagement, which vary in each community. In 2017, Winangali (an Indigenous-owned engagement consultancy) evaluated the RSAS in partnership with commissioned evaluator, Ipsos. They chose a realist-informed, participatory and qualitative approach on the basis that this would be culturally appropriate, and would recognise context, complexity and diversity among families and communities. The evaluation involved:

- developing multiple theories of change on how families may respond to the RSAS, drawing from the literature, program data, anecdotal evidence and previous evaluations
- using local Aboriginal and Torres Strait Islander field researchers, and working with communities to identify a suitable researcher
- conducting 114 semi-structured interviews with parents and carers, and group discussions with parents, carers, local community members and RSAS staff
- researchers participating in a workshop to identify key themes, patterns and points of contention that arose during the fieldwork to develop and refine their understanding of how the initiative was working.

The evaluation found that the RSAS did not improve school attendance for all families and when acting in isolation it was unable to improve school attendance for families experiencing complex life events. However, the evaluation also found that the RSAS was an effective referral and coordination point, but only when staff were appropriately trained, knowledgeable about other services, and were seen as ‘the right person for the job’.

The evaluation identified four different family types served by the RSAS, each requiring different kinds of support:

- *committed families* — who see school as important and a stepping stone to higher education and jobs, but occasionally need practical support to get their children to school
- *protective families* — who want their children to attend school, but worry that school (or travel to school) is unsafe and so protect their children by keeping them at home
- *unsure families* — who are not convinced that education would lead to opportunities for their children, and worry that schools would not teach or value their culture
- *disconnected families* — who want their children to go to school, but are isolated and lack community ties, and do not know how to get children to school or who to ask for help.

The evaluation also found that different types of families responded to the same activity in different ways, and that a critical factor for the effectiveness of the RSAS was the ability of staff to recognise and respond to the different needs of families. As a result, program staff received additional training on the different family types and guidance on how to best support families.

Sources: ANAO (2019b); DPMC (2018d).

Both developmental and realist approaches are useful in situations where the relationship between inputs and effects are non-linear, where multiple, interacting variables influence an outcome, and in cases of uncertainty and rapid change. However, while these approaches can provide insights on how, and why a program works, or does not work, they do not

measure overall impact, causality or whether change has occurred compared to other scenarios. Findings are also context-specific and are not generalisable.

Ethical practice for quality evaluations

Applying ethical standards improves the quality and consistency of evaluation and ensures that being engaged in an evaluation has a positive impact on Aboriginal and Torres Strait Islander people. All stages of evaluation — planning, commissioning, conduct, reporting and use — should be conducted in an ethical way.

Ethical practice during evaluation should be guided by existing ethical guidelines for research involving Aboriginal and Torres Strait Islander people, including those developed by AIATSIS and the National Health and Medical Research Council.³

Formal review by an ethics committee will be required for some, but not all, evaluations conducted under the Strategy. For example, where an evaluation only uses existing or de-identified data, formal ethical review may not be necessary. However, all evaluation projects should include a systematic and well-documented assessment of ethical risks.

Evaluation budgets and timeframes need to ensure evaluators can 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 capacity and report back to evaluation participants.

Evaluation capability

Evaluation capability, including cultural capability, is also important for high quality evaluations and the validity of evaluation 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 an agency. Agency staff need to be able to plan for evaluations, collect data and interpret and implement findings. They also need to be able to commission an evaluation appropriately, including selecting an evaluator, manage an evaluation and assess its quality.

Evaluators need the skills to conduct evaluations well. This means being able to apply systematic methods of inquiry that are appropriate to the evaluation in question to ensure credible results. It also means undertaking culturally capable and ethical evaluation (which is critical for valid results, box 8). As BetterEvaluation said,

Without recognizing the influence of culture on human behaviour, the evaluators may arrive at conclusions which are fundamentally flawed. (2016)

³ See AIATSIS (2019) and NHMRC (2018a).

Box 8 What is culturally capable evaluation?

Cultural capability is central to the validity of evaluative conduct and conclusions. Cultural capability is a reflective way of doing and thinking (or an attitude towards culture) rather than particular skills or knowledge. It goes beyond conducting evaluations in culturally appropriate ways, rather it should draw on the priorities, perspectives and values of Aboriginal and Torres Strait Islander people. This is so that definitions of ‘good program content and design’, ‘high quality implementation and delivery’ and ‘outcomes of value’ reflect what Aboriginal and Torres Strait Islander people value (ANZEA 2011, p. 10).

Evaluation teams for policies and programs affecting Aboriginal and Torres Strait Islander people should have skills and experience in working with Aboriginal and Torres Strait Islander people. There are a number of ways to improve cultural capability including:

- evaluators and commissioners working with Aboriginal and Torres Strait Islander people when developing and assessing evaluation methodologies
- employing Aboriginal and Torres Strait Islander people as commissioners and/or evaluators
- working with Aboriginal and Torres Strait Islander consultants who can advise on cultural competency or safety.

The Strategy’s principles-based guidance on how to conduct evaluations is expected to increase demand for culturally capable evaluators. There should also be increased opportunities for Aboriginal and Torres Strait Islander people to be engaged in, participate in, and provide advice on evaluation design, data collection and the interpretation of evaluation results. This in turn should strengthen the capability of Aboriginal and Torres Strait Islander people to undertake evaluations.

The Head of Profession (agreed to as part of the Australian Government’s response to the Independent Review of the APS to develop core competencies and establish learning and development frameworks for relevant professions), should be well placed to support agencies to build capability. It could:

- provide training for commissioning, conducting, and using evaluations of policies and programs that affect Aboriginal and Torres Strait Islander people
- facilitate an APS wide community of practice for people who are involved in evaluating policies and programs that affect Aboriginal and Torres Strait Islander people
- establish processes through which evaluators can seek secondments or other mobility opportunities, with a view to broadening or deepening their experiences evaluating policies and programs affecting Aboriginal and Torres Strait Islander people
- develop a strategy to build a cohort of Aboriginal and Torres Strait Islander evaluators within the APS (this could include structured training, networking, mentoring and secondment opportunities).

Quality assurance and external oversight to improve the quality of evaluations

Evaluations should have quality assurance processes in place (such as peer review). External oversight of evaluations can also improve the quality of evaluations. By increasing accountability, external oversight provides an incentive for agencies to ensure that the evaluations they undertake are of high quality, are ethical, are useful, and that their evaluation processes, practices and outputs align with the Strategy's principles.

While agencies have some incentive to evaluate their policies and programs, we found this is often more about accountability and compliance than finding out whether or not policies or programs are working, or how they can be improved. Without external oversight, there is a risk that the Strategy will not change how agencies approach evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, or that some agencies will follow the Strategy's guidance but not others.

The importance of an independent oversight body for the Strategy to be effective was acknowledged by a number of participants. Some supported the Productivity Commission having this role,⁴ with some proposing that the Commission be augmented by Aboriginal and Torres Strait Islander representation or an Indigenous Council. Others proposed that the oversight body be Indigenous-led, or for the oversight role to be allocated to a non-government Aboriginal and Torres Strait Islander body, while others suggested that the oversight role could sit with AIATSIS, the Australian National Audit Office or the Australian Human Rights Commission.

Independence for the Strategy's monitoring arrangements is critical for credibility. The above-mentioned statutory bodies have the independence from the day-to-day activity of government necessary to objectively assess agencies' performance against the Strategy. However, such bodies do not have powers to direct government agencies in the exercise of their day-to-day functions. It would be transparency of agencies' performance that would provide the incentive for agencies to lift the bar on the quality and usefulness of evaluations.

But while external oversight is important for lifting the quality of evaluation, it is just one of the central evaluation functions needed for an effective Strategy and to support a culture of evaluation. Providing evaluation leadership (including sharing innovations in evaluation and engagement practices), providing advice to agencies on what the Strategy means in practice, and identifying high priority and cross-cutting evaluations, are also important central evaluation functions.

And ideally, these central evaluation functions would be located in one body. It is on this basis that we are proposing that a dedicated unit — an Office of Indigenous Policy Evaluation (OIPE) — be established within an existing independent statutory authority

⁴ The Letter of Direction for the Strategy implies that the Productivity Commission will play an oversight role for the Strategy. It states that 'the Commission will review the performance of agencies against the Strategy over time, focusing on potential improvements and on lessons that may have broader application for all governments'.

(views are being sought on which existing authority) to oversee the implementation of the Strategy and coordinate a whole-of-government approach to evaluating policies and programs affecting Aboriginal and Torres Strait Islander people. The OIPE would:

- have stewardship of the Strategy
- champion the value of evaluation in policy making and identify potential areas for evaluations or systematic reviews on cross-cutting topics
- advise Australian Government agencies on what the Strategy means in practice
- identify evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people
- oversee agencies' Three Year Evaluation Forward Work Plans
- coordinate with the Head of Evaluation Profession on capability building opportunities
- monitor and publicly report on Australian Government agencies' performance against the Strategy, and on how effective the Strategy has been in encouraging high quality and useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

An important role for the OIPE would be to encourage agencies to evaluate with the purpose of improving policy and program outcomes for Aboriginal and Torres Strait Islander people. This includes experimentation and trialling innovations (noting that the answers to many policy problems are not known in this area). It also means engaging with Aboriginal and Torres Strait Islander people in such a way that meaningful information is collected, that is, information that helps policy makers better understand the impacts of policies and programs.

As mentioned earlier, in line with the principle of the Strategy of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges, we are also proposing that an Indigenous Evaluation Council (the Council), with a majority of Aboriginal and Torres Strait Islander members, be established. The Council would provide strategic guidance to the OIPE on its stewardship and monitoring of the Strategy as well as partner with the OIPE to identify whole-of-Australian Government evaluation priorities (see also below the section on a whole-of-government approach). The Council would also partner with the OIPE to report on the state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people across the APS.

Enhancing the use of evaluation evidence

Evaluation is a key tool for evidence-based policy. But for evaluation findings to be used, they need to be relevant and timely for decision making. Evaluations should be planned, commissioned and conducted with a clear intention to use the resulting analysis and findings.

The use of evaluation findings can be enhanced throughout the evaluation process by evaluators asking: what should be evaluated, for what purpose, when, how and with whom? The World Bank Group talks about '... [optimising] the value of evaluations by adopting

an evaluation use perspective throughout the entire evaluation process’ (2019, p. 10). A *Guide to Evaluation under the Indigenous Evaluation Strategy*, which accompanies the Strategy sets out the questions that evaluators should be thinking about at the various stages of the evaluation process to improve the usefulness of evaluations.

The usability of evaluation findings will also be enhanced if evaluators engage with those affected by the policy or program early in the evaluation process to find out what information and decision-making needs such groups have (which should include reflecting on what to evaluate and the questions the evaluation should focus on) and what timing is most appropriate for influencing decisions about the policy or program.

Requirements under the Strategy for New Policy Proposals for policies and programs (those that will have a significant impact on Aboriginal and Torres Strait Islander people) to have an evaluation plan, and Australian Government agencies to have an annual rolling Three Year Evaluation Forward Work Plan, will help embed evaluation into policy planning and policy making processes.

Disseminating evaluation results

Publication of evaluation results is important for improving policies and programs — policy makers need to know about evaluation evidence to use it. Publication can also improve the quality of evaluations (by allowing independent scrutiny), facilitate more informed public debate about the impact of policies and programs, and enhance accountability by increasing visibility and pressure for agencies to follow up with a management response to evaluation findings. Evaluation results are also more likely to be used if they are synthesised into the existing evidence base and tailored for specific users.

But under current arrangements, it is not easy for policy makers and other evaluation users to access evaluation evidence (nor is it easy to identify evidence gaps that evaluations could fill). Many evaluations are not published, and evaluation evidence is not collected in one central place.

Australian Government agencies should publish all evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. Where there are concerns that publishing a full evaluation report would compromise confidentiality or privacy, or where there is culturally sensitive information, we are recommending a summary report be published instead. All published evaluation reports should have a clear and concise summary of the evaluation findings. Evaluation reports should document details of data collected, approaches and methods used, ethical practices, limitations of the evaluation, and costs. The evaluation summary report should also document how the evaluation adhered to the principles of the Strategy.

Consistent with good ethical research practice, agencies should share evaluation findings with Aboriginal and Torres Strait Islander people, communities and organisations who are participants in evaluations or stakeholders in the policies and programs being evaluated.

Agencies should also publish their management response to evaluation findings (this could be included as part of the evaluation report or separately). This 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.

A central clearinghouse should be established to maximise the accessibility of evaluation reports, but also to synthesise evaluation findings into the broader evidence base and to translate those conclusions into accessible products for different audiences. This will help to support wider, better and more informed use of evaluation evidence, not only by policy makers but also by researchers and service providers. It is rare that a single evaluation provides sufficient robust information to generalise into conclusive lessons about a policy or program. Rather, it is the accumulation of consistent evidence over time and across different contexts that gives policy makers confidence in evaluation results.

An evaluation culture that values evidence and innovation

Strong leadership is important for building an evaluation culture that values evidence and innovation, shares knowledge (including mistakes) and learns from the results of evaluations. The OIPE would play an important role in championing the value of evaluation — helping agencies to view evaluation less as a compliance exercise, and more as a key tool for finding innovative solutions to complex policy problems. The transparent sharing of lessons, including lessons that have broader application for evaluation of policies and programs, will also be important for creating a culture that values evaluation and innovation.

Promoting a whole-of-government approach to priority setting

As discussed earlier, currently there is no strategic whole-of-government approach for prioritising evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. This means Australian Government agencies use different approaches for setting evaluation priorities. There is also little transparency or accountability around how agencies make decisions about what to evaluate. And while Aboriginal and Torres Strait Islander people's perspectives on what policies and outcomes matter most are critical for identifying priorities for evaluation, engagement with Aboriginal and Torres Strait Islander people on the priority areas for evaluation is limited.

To help guide agencies' efforts on what to evaluate under the Strategy, the Australian Government should commit to establishing a set of priority evaluation areas — which should be decided on in partnership with Aboriginal and Torres Strait Islander people.

The Commission is proposing that the interim set of government-wide priority areas for evaluation under the Strategy be based on the initial outcomes from the Joint Council on Closing the Gap (a formal partnership arrangement between governments and representatives of Aboriginal and Torres Strait Islander people). Between the draft and final report, we will engage further and refine the interim set of priorities.

The priorities will be formalised as part of the monitoring process proposed for the Strategy, working in partnership with the Council. It is expected that evaluation priorities — including opportunities for cross-agency evaluations — will continue to be better defined over time with better information on:

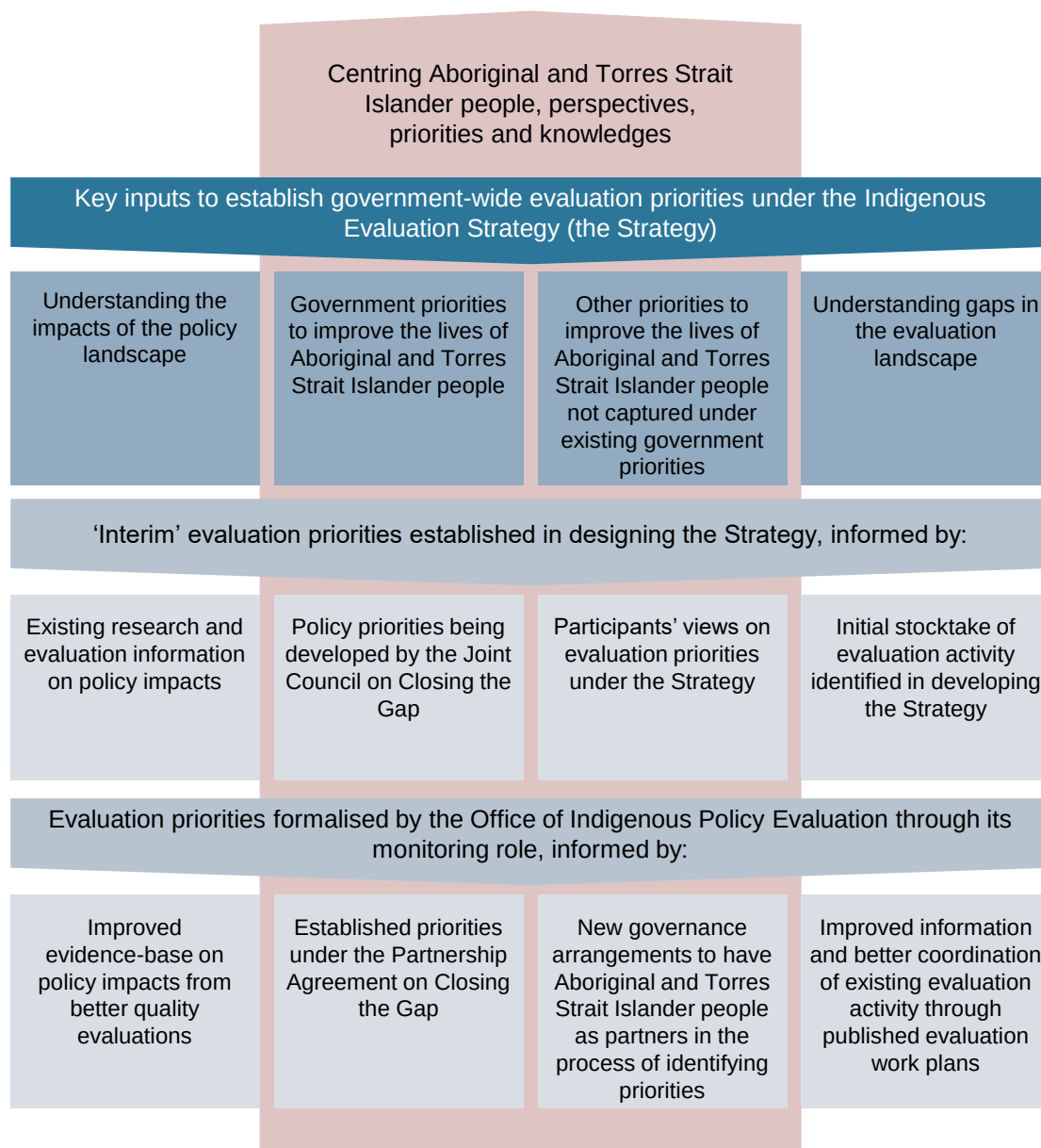
- the policy and reform priorities agreed to by the Joint Council on Closing the Gap
- the evaluation and policy landscape of agencies as they report more and better quality information through forward work plans and completed evaluation reports on key policies and programs
- other evaluation priorities identified through consultations as part of this monitoring process (figure 7).

These centralised evaluation priorities should not diminish the responsibilities agencies have for assessing how their policies and programs impact on Aboriginal and Torres Strait Islander people and communities, or on the state of the evidence base for their key policies and programs. Agencies should identify new and existing policies and programs that contribute to these government-wide priority areas, and assess whether they should be prioritised for evaluation.

Under the Strategy, agencies will be required to publish an annual rolling Three Year Evaluation Forward Work Plan, which details what evaluations are to be undertaken over the coming period (and when), how the evaluations were selected, and how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges were centred as part of the prioritisation process. While evaluation priorities will be identified for three years going forward, there will also be flexibility to respond to changing circumstances.

Monitoring of the Strategy will include an assessment of whether agencies have implemented their evaluation plans and the extent to which high priority evaluations are taking place.

Figure 7 **Process for formalising government-wide evaluation priorities**



4 Review and future governance arrangements

Reviewing the Strategy

As discussed above, as part of monitoring the Strategy's implementation, the Commission proposes that the OIPE would monitor the performance of the Strategy. However, along with the monitoring process, it is also important that a separate process for reviewing the Strategy be established. This review process should assess:

-
- whether the Strategy’s principles remain fit-for-purpose
 - the extent that the Strategy has been effective in encouraging higher quality and more useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, including the effectiveness of the Strategy’s actions
 - the extent that the performance of those tasked with overseeing the Strategy has been effective and whether changes to the policy environment require changes to oversight arrangements.

A Centre for Evaluation Excellence

The case for central evaluation leadership and external oversight is wider than the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. As such, over the longer term (five years), the Commission is proposing that a new independent Australian Government agency — a Centre for Evaluation Excellence (CEE) — be established.

The CEE would have similar leadership and oversight roles as the OIPE (figure 8), however, it would lead on efforts to improve evaluations on social and health policies across the APS as a whole. The Commission also proposes that when the CEE is established, the OIPE would move over to the CEE and operate as a standalone branch in its leadership of the Indigenous Evaluation Strategy (with the Indigenous Evaluation Council continuing its role providing strategic direction and partnering with the OIPE’s work).

The OIPE would supplement the functions of the CEE by providing specialised advice and guidance on issues specific to the Indigenous Evaluation Strategy (including how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred throughout the evaluation process).

Figure 8 **Functions of the Centre for Evaluation Excellence**

Centre for Evaluation Excellence	
	Office of Indigenous Policy Evaluation
Leadership/advocacy: Evaluation champion	Evaluation champion
Capacity building: • Guidance material and advice to agencies on good practice evaluation governance and standards • Collaborate with Head of Evaluation Profession (on skill development, communities of practice, etc.)	• Guidance and advice on Indigenous Evaluation Strategy governance and standards, and good practice • Cultural capability (so evaluations centre Indigenous people, perspectives, priorities and knowledges) • Collaborate with APS Head of Evaluation Profession
Coordinate evaluation effort: Advise on evaluation priority-setting if needed	Government-wide priority-setting under the Strategy
Oversight: Agency evaluation plans	Oversee agency Three-Year Forward Work Plans under the Strategy
Clearinghouse for evaluation reports: Translate and disseminate learnings	Translate and disseminate learnings under the Strategy
Monitor and report: Whole-of-government evaluation performance	Monitor and report on performance under the Strategy

Information requests

In addition to the Commission seeking feedback on the *Indigenous Evaluation Strategy, Draft* and *A Guide to Evaluation under the Indigenous Evaluation Strategy*, there are some areas in this background paper where more information would assist in further developing the Strategy. These information requests are collated below.

Chapter 5 — Evaluating for quality results

Governments across Australia are recognising the benefits of genuine engagement and co-design with Aboriginal and Torres Strait Islander people. However more information is needed to understand how co-design works in practice, both in policy and program design, and in evaluation itself. The Commission is interested in hearing participants' views on what effective engagement models for policy and program evaluation look like.

INFORMATION REQUEST 5.1

The Commission is seeking information on effective engagement strategies for evaluation. What engagement models are most effective? For what types of evaluations is co-design most useful? Why?

Chapter 6 — What to evaluate

The Commission is proposing a set of 'interim' government-wide evaluation priorities based on the Council of Australian Governments and the Joint Council on Closing the Gap draft policy priorities to help focus agency evaluation efforts under the Strategy.

INFORMATION REQUEST 6.1

The Commission proposes that the interim evaluation priorities for the Indigenous Evaluation Strategy be broad and based on the Council of Australian Governments and the Joint Council on Closing the Gap draft policy priorities. Is this appropriate?

If so, are there any priority areas missing from those currently identified through the Council of Australian Governments and the Joint Council on Closing the Gap?

If not, what specific policy or program areas should be the focus of better quality evaluations?

Chapter 7 — Developing a culture of evaluation and learning

The Commission is proposing that an Indigenous Evaluation Clearinghouse for evaluations conducted under the on Strategy be established within an existing independent statutory authority. It is proposed that the Clearinghouse would collect and publish evaluation reports or summaries online; disseminate evaluation findings to a variety of audiences such that they can best be put into practice; and synthesise new evaluation findings into the existing evidence base.

INFORMATION REQUEST 7.1

The Commission is seeking participants' views on which current Australian Government agency would be best placed to house the Indigenous Evaluation Clearinghouse: the Australian Institute of Health and Welfare; the Australian Institute of Family Studies; the Australian Institute of Aboriginal and Torres Strait Islander Studies; the Productivity Commission; or some other agency?

The Commission is proposing that a central advisory and oversight function under the Strategy — an Office of Indigenous Policy Evaluation (OIPE) — be established as a standalone unit within an existing independent Australian Government statutory authority.

INFORMATION REQUEST 7.2

The Commission is seeking participants' views on the location of an Office of Indigenous Policy Evaluation to oversee Australian Government agencies' implementation of the Indigenous Evaluation Strategy. Which current statutory agency would provide the best location for the Office of Indigenous Policy Evaluation, and why?

The Commission is proposing that an Indigenous Evaluation Council ('the Council') be established to provide strategic guidance to the OIPE on its advisory and oversight work. The Council would also partner with the OIPE to identify government-wide evaluation priorities under the Strategy. In line with the principle of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges, the Commission is proposing that the majority of Council members be Aboriginal and Torres Strait Islander people.

INFORMATION REQUEST 7.3

The Commission is seeking participants' views on how members of an Indigenous Evaluation Council might be appointed. For example, could members be nominated by the Aboriginal and Torres Strait Islander community controlled sector; for their experience in research, monitoring and evaluation; or based on some other factors? Would the host agency for the Office of Indigenous Policy Evaluation, and/or the Australian Government, need to be members?

Chapter 9 — Data for evaluation

The Commission is proposing the Australian Institute of Health and Welfare work in partnership with Aboriginal and Torres Strait Islander people to develop and validate a data dictionary containing a set of core indicators of Aboriginal and Torres Strait Islander people's wellbeing that could be used for evaluations. Ideally, the data dictionary would: ensure greater consistency and comparability across government; minimise duplication and differing definitions; and reduce the costs and burdens of developing data collection tools both for agencies and Aboriginal and Torres Strait Islander people.

INFORMATION REQUEST 9.1

The Commission is seeking feedback on the indicators that are highest priority for inclusion in the data dictionary. What outcomes and indicators are most informative for government and Aboriginal and Torres Strait Islander people?

Data linkage can be a powerful tool for use in evaluation by bringing together multiple datasets to gain insights that are not possible with a single dataset or survey. There are potentially many linked datasets that could be useful for evaluations under the Strategy.

INFORMATION REQUEST 9.2

Which data linkages would best support the Indigenous Evaluation Strategy? Which data linkages are important to Aboriginal and Torres Strait Islander people?

1 About the Strategy

The Australian Government delivers a wide range of policies and programs to Aboriginal and Torres Strait Islander people. It is critical that these policies and programs are evaluated, because the information generated by evaluations has the potential to improve policies and programs (by answering questions about effectiveness, value for money and relevance), and in turn, improve the lives of Aboriginal and Torres Strait Islander people.

But for evaluation to influence policies, it needs to be high quality and useful. And in the context of policies and programs affecting Aboriginal and Torres Strait Islander people, this means taking into account the priorities, world views and lived experiences of Aboriginal and Torres Strait Islander people. This is a shift from the way evaluations of many policies and programs affecting Aboriginal and Torres Strait Islander people are currently undertaken. Maggie Walter, for example, said:

Currently, evaluations only superficially reflect Aboriginal and Torres Strait Islander people's perspectives because input is only sought at the participant level. By that time, important decisions have already been made about the evaluation process — what to evaluate, how to evaluate, the evaluation logic, key outcomes and objectives. Indigenous voice is denied and the current policy mindset is perpetuated. We need Indigenous voice up-front and helping to make decisions from the outset. (sub. 112, p. 2)

The key to good policy outcomes is putting the people who are intended to benefit from the policy at the centre of policy development and evaluation. Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges should be reflected in what is evaluated, how evaluation is undertaken, and the outcomes policies seek to achieve.

1.1 Some background to the Strategy

For decades there have been calls for better evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. Back in 2009, the Department of Finance and Deregulation's *Strategic Review of Indigenous Expenditure* argued that a more rigorous approach to program evaluation was needed.

Robust evidence is lacking on the performance and effectiveness of many Indigenous programs. Program evaluation activity in this area has been patchy at best, and many of the evaluations which have been conducted have lacked a suitable measure of rigour and independence. More robust evaluation arrangements are needed for the future. (2010, p. 12)

Participants of the Productivity Commission's (the Commission) 2012 roundtable, *Better Indigenous Policies: The Role of Evaluation*, also concluded that there were significant gaps

in the evidence base for policies and programs affecting Aboriginal and Torres Strait Islander people and that an overarching policy review was required to address the way governments work with, and in, Indigenous communities (PC 2013). And one of the key messages of the Commission's *National Indigenous Reform Agreement Performance Assessment 2013-14* was that there was extensive reporting on outcomes for Aboriginal and Torres Strait Islander people, but more attention and resources needed to be given to policy evaluation (PC 2015).

In the 2017-18 Budget, new measures for research and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people were announced, including:

- \$40 million over four years to strengthen evaluation of the Indigenous Advancement Strategy (IAS, box 1.1)
- \$10 million over three years to establish an Indigenous Policy Research Fund to add to the Indigenous policy evidence base
- \$2.9 million over four years for the Productivity Commission to enhance its role in Indigenous policy evaluation and for an additional Commissioner with relevant experience in Indigenous policy. It was also announced that the Commission would develop an Indigenous Evaluation Strategy (the Strategy) to be reported against by all Australian Government agencies (DPMC 2017a).

Box 1.1 About the Indigenous Advancement Strategy

The Indigenous Advancement Strategy (IAS) is one of the ways the Australian Government funds and delivers programs designed specifically for Aboriginal and Torres Strait Islander people. The IAS reflects the Australian Government's objective of achieving better results for Aboriginal and Torres Strait Islander people in three priority areas — getting children to school, adults into work and building safe communities. The IAS consolidates the many different Indigenous policies and programs that were delivered by the Australian Government into five overarching programs: Jobs, Land and Economy; Children and Schooling; Safety and Wellbeing; Culture and Capability; and Remote Australia Strategies.

In the 2019-20 Budget, the Australian Government allocated \$5.2 billion to the IAS over four years.

Source: NIAA (2020).

More about what the Commission has been asked to do

The Commission has been asked to develop a whole-of-government (Australian Government) evaluation strategy for policies and programs affecting Aboriginal and Torres Strait Islander people.

Improving outcomes for Indigenous Australians depends on agencies with responsibility for policies and programs affecting Indigenous Australians undertaking meaningful evaluations. The Commission is to develop a strategy to guide that evaluation effort. (Frydenberg 2019)

The evaluation strategy is to include:

- a principles-based framework for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people
- priorities for evaluation
- an approach for reviewing agencies' conduct of evaluations against the strategy.

The Letter of Direction states that the Commission will review the performance of agencies against the strategy over time, focusing on potential improvements and lessons that may have broader application for all government. The Commission is also to consider areas where there may be value in it undertaking evaluation and to provide a forward work program.

1.2 What is evaluation and why evaluate?

Policies and programs are put in place by governments to improve outcomes for individuals and the community, but governments can never be certain how the policies will play out in practice (particularly because the way policies and programs are implemented can affect intended benefits and program success). Policies and programs can also have unintended consequences. And there are costs to consider — a program could achieve its intended outcomes, but at a considerably higher cost than expected. As Rossi, Lipsey and Henry put it:

... there are many ways for a program to fail to produce the intended benefits ... Good intentions and plausible program concepts are not sufficient. If they were, we could be confident that most social programs are effective at delivering the expected benefits without conducting any evaluations of their theories of action, quality of implementation, positive and adverse effects, of benefit-cost relationships. Unfortunately, that is not the world we live in. ... Assessing the effectiveness of social programs and identifying the factors that drive or undermine their effectiveness are the tasks of program evaluation. (2018, p. 4)

Evaluation is the systematic process of making a judgment about the merit or worth of a policy or program. It is a key tool for evidence-based policy. Reasons for undertaking evaluations include to:

- better understand what policies and programs work and why (and when and for whom) — to support evidence-informed policy development, program design and implementation
- inform improvements in policies and programs — to support learning by doing
- inform policy makers about how to best use taxpayers' money (shedding light on whether program funds are properly spent, are delivering value for money and are meeting recipients' needs) — to support accountability.

The evidence from evaluation can also improve the community's understanding about policy and program effectiveness which can lead to deeper and more mature policy debates. George Argyrous, from the Australia and New Zealand School of Government's (ANZSOG's) Evidence and Evaluation Hub, said:

... good evaluations have debunked, or at least queried, the credential of many programs that were originally accepted without question. (ANZSOG 2018, p. 1)

Evaluation can also help to build trust, particularly if findings are used to support 'learning by doing' and stakeholders work together to find ways to achieve agreed outcomes.

A number of participants to this project identified evaluation (when conducted and used well), as a key tool to improving policies and programs and, in turn, the lives of Aboriginal and Torres Strait Islander people (box 1.2). Others pointed to some of the consequences of not evaluating or of undertaking poor quality or poorly focused evaluations — ad hoc policy decision making, not knowing whether the program or policy is achieving anything (or whether it is doing harm), and the indiscriminate cutting of policies and programs.

The importance of evaluation holding governments accountable for outcomes, and identifying opportunities to use funds more effectively, was also highlighted (box 1.2).

1.3 Why a whole-of-government Indigenous evaluation strategy?

A better evidence base is needed to answer questions about the effectiveness (or otherwise) of policies and programs affecting Aboriginal and Torres Strait Islander people. Despite the calls for better quality evaluations and more evidence-informed policies, 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 (ANAO 2019b; Empowered Communities 2015; Hudson 2017; Muir and Dean 2017). Empowered Communities, for example, called for greater insights on policy and program effectiveness:

We need to know what is working and if policy and programs which apply to Indigenous individuals, families and communities are having an impact. We also want greater insight into

why policy or program implementation is not effective, and we need early opportunities for correction or reinvestment of funds and effort to ensure that funding is directed to where it is needed most. (sub. 41, p. 5)

Box 1.2 Participants saw evaluation as a key tool for achieving better outcomes

Indigenous Business Australia:

... evaluation is an essential part of policy development and program management. By providing a process through which the goal of policy frameworks and programs can be questioned and how they are performing and what is working and what is not, evaluation can provide policy and program stakeholders and designers with valuable information about performance. (sub. 36, p. 4)

Chelsea Bond et al.:

In general terms, we agree that evaluation is important for guiding evidence-informed programming to alleviate social disadvantage. It is our experience that Indigenous community-controlled sector ... organisations ... and peoples want evidence that the policies and programs they are delivering, or participating in, are effecting positive change. They also want to see government and mainstream agencies face the same level of scrutiny, reporting and evaluation, and be held to the same high standards. (sub. 40, p. 3)

Queensland Aboriginal and Islander Health Council:

There is value in a process that critically examines a program, by collecting and analysing information about a program's activities, characteristics, and outcomes. ... Consultative, structured and reciprocated evaluation, ensures that the needs of Aboriginal and Torres Strait Islander peoples and programs are met. If used responsibly and respectfully, evaluation findings can inform best practice and strengthen continuous quality improvement. (sub. 55, p. 6)

Indigenous Community Volunteers:

Public expenditure on government programs should be informed by, and reported on, using good practice evaluation. (sub. 88, p. 2)

Minerals Council of Australia:

Enhancing evaluation practices will assist to improve shared understanding of partnership outcomes and provide strong evidence base for future collaborations. (sub. 34, p. 1)

The Centre for Independent Studies, in a report on *Mapping the Indigenous Program and Funding Maze*, also said:

If Indigenous people are ever to benefit from the considerable public and private investment into Indigenous affairs, the government has to cease funding and delivering programs without any evidence of their efficacy and without any consultation with Indigenous communities. (Hudson 2016, p. 28)

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 (box 1.3).

Box 1.3 **Wide gaps remain after more than ten years of Closing the Gap initiatives**

Despite more than ten years of Closing the Gap initiatives, wide gaps between Aboriginal and Torres Strait Islander people and non-Indigenous people persist across many indicators of wellbeing. Four of the seven Closing the Gap targets have expired without being met. Just two of the ongoing targets, early childhood education and Year 12 attainment, are on track.

- The target to halve the gap in child mortality rates by 2018 was **not met**. Since the target baseline (2008), Indigenous child mortality rates have declined more slowly than non-Indigenous child mortality rates.
- The target to have 95 per cent of Indigenous four year olds enrolled in early childhood education by 2025 is **on track**. In 2018, 86 per cent of Indigenous children were enrolled in early childhood education programs, which was higher than the agreed trajectory point for 2018 to reach the target by 2025.
- The target to close the gap in school attendance by 2018 was **not met**. Attendance rates for Indigenous students did not improve between 2014 and 2019 (about 82 per cent in 2019) and remain below the rate for non-Indigenous students (about 92 per cent).
- The target to halve the gap in reading and numeracy by 2018 was **not met**, with about one in four Indigenous students in Years 5, 7 and 9, and one in five in Year 3, remaining below national minimum standards in reading in 2018, and between 17 to 19 per cent of Indigenous students being below the national minimum standards in numeracy.
- The target to close the gap in life expectancy by 2031 is **not on track**. The 2020 Closing the Gap Report indicated that between 2006 and 2018, there was an improvement of almost 10 per cent in Indigenous age-standardised mortality rates, although non-Indigenous mortality rates improved at a similar rate.
- The target to halve the gap in Year 12 attainment or equivalent by 2020 is **on track**. Between 2008 and 2018-19, the proportion of Indigenous Australians aged 20–24 years attaining Year 12 or equivalent increased by about 21 percentage points. The gap narrowed by about 15 percentage points, with non-Indigenous attainment rates improving at a slower pace.
- The target to halve the gap in employment by 2018 was **not met**, with no significant improvement in the gap. The employment rate for Indigenous Australians increased by 0.9 percentage points, while falling 0.4 percentage points for non-Indigenous Australians.

Sources: AIHW (2017b); Australian Government (2020b); DPMC (2019a).

Another concern is the quality of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people (chapter 3). The Centre for Independent Studies (Hudson 2017, p. 13) reviewed 49 Indigenous-specific programs and concluded that ‘in general, Indigenous evaluations are characterised by a lack of data and the absence of a control group [necessary for *some* evaluation methods, chapter 4], as well as an over-reliance on anecdotal evidence’.

An Indigenous Evaluation Strategy is an opportunity to improve the quality and use of evaluations of government policies and programs affecting Aboriginal and Torres Strait Islander people, and to provide insights more generally about how to assess complex programs affecting all Australians. A Strategy can also provide a whole-of-government

approach to evaluation priority-setting and the conduct of evaluations (for a more consistent approach across government agencies).

And importantly, and in keeping with a commitment by the Australian Government to work differently with Aboriginal and Torres Strait Islander people (box 1.4), the Strategy is a critical opportunity to revisit how Australian Government agencies engage and work with Aboriginal and Torres Strait Islander people, communities and organisations when evaluating policies and programs. A number of participants saw the Strategy as a way to strengthen Aboriginal and Torres Strait Islander people's capability, leadership and decision making in evaluation (box 1.5).

The Strategy also has the potential to shift Aboriginal and Torres Strait Islander people's views about the value of evaluation. The Commission heard repeatedly that Aboriginal and Torres Strait Islander people were often the subject of research and evaluation, but they rarely saw the benefits of such work. The National Health and Medical Research Council also recently said:

... the early colonisers knew nothing about Aboriginal and Torres Strait Islander Peoples and their cultures. The substantial errors of judgement and the misconceptions that followed have had a significant impact on Aboriginal and Torres Strait Islander Peoples ever since. Aboriginal and Torres Strait Islander Peoples are considered to be the most researched peoples in the world. Despite this, [Aboriginal and Torres Strait Islander Peoples] have received very little perceivable benefit. (2020, p. 1)

More recently, an increase in the number of Aboriginal and Torres Strait Islander organisations, evaluators and research bodies, and their stewardship and advocacy for evaluation to improve lives — has seen a shift in the views about evaluation of many Aboriginal and Torres Strait Islander people. The Indigenous Evaluation Strategy should add further impetus to these changing views.

The Strategy is for Australian Government agencies, but it could also have benefits for state, territory and local governments. As the Western Australian Government (sub. 74) suggested, the Strategy would have implications for state and territory governments given the significant crossover between Australian and State and Territory government responsibility.

The Minerals Council of Australia, while acknowledging that the Strategy is intended for Australian Government agencies, suggested that a principles-based strategy could also be adopted for a range of evaluation purposes, including partnerships between private sector companies and Aboriginal and Torres Strait Islander organisations, and to assist multi-sector collaborations to support the priorities of Aboriginal and Torres Strait Islander people (sub. 34).

Box 1.4 **A changing policy landscape**

In March 2019, a Partnership Agreement between the Council of Australian Governments (COAG) and Aboriginal and Torres Strait Islander people, through the National Coalition of Aboriginal and Torres Strait Islander Peak Organisations (Coalition of Peaks), came into effect. The Agreement sets out how governments and the Coalition of Peaks will work together over the next ten years on Closing the Gap.

The Joint Council on Closing the Gap is the first ever COAG Council to include members from outside governments, through membership of representatives of the Coalition of Peaks. The historic Agreement ‘embodies the belief of all its signatories that shared decision making with Aboriginal and Torres Strait Islander community controlled representatives in the design, implementation and monitoring of the Closing the Gap framework is essential to achieve their shared goal to close the gap in life outcomes between Indigenous and non-Indigenous Australians’ (JCOCTG 2019c, p. 2). The Joint Council has determined priority reform areas focusing on greater shared decision making, building the community controlled service sector, and ensuring mainstream government agencies and institutions undertake systemic and structural transformations to contribute to Closing the Gap.

Aboriginal and Torres Strait Islander self-determination is recognised as key to achieving changes in the lives of Aboriginal and Torres Strait Islander people and the Agreement embodies that intent.

The Australian Government also recently committed \$7.3 million for a co-design process to develop options for a ‘voice to government’ for Aboriginal and Torres Strait Islander peoples and a model to improve local and regional decision making (Wyatt 2019). When announcing the co-design process, the Minister for Indigenous Australians said:

The best outcomes are achieved when Indigenous Australians are at the centre of decision-making. ... It’s time that all governments took better steps to empower individuals and communities, and work in partnership to develop practical and long lasting programmes and policies that both address the needs of Indigenous Australians and ensure that Indigenous voices are heard as equally as any other Australian voice. (Wyatt 2019)

The 2019 Independent Review of the APS also recommended that the Australian Government and APS work as genuine partners with Aboriginal and Torres Strait Islander peoples, including through the application of place-based decision-making frameworks and a recognition of ‘the importance of Aboriginal and Torres Strait Islander peoples making decisions on matters affecting their lives and communities’ (DPMC 2019e, p. 132).

The Australian Government agreed to the APS applying the recently agreed APS Framework for Engagement and Participation to ‘support genuine collaborations with Australians in designing better services and finding solutions to policy problems’ (DPMC 2019d, p. 17).

Sources: COAG (2018); DPMC (2019a); JCOGTG (2019a); Morrison (2020); Wyatt (2019).

Box 1.5 Participants identified benefits of a Strategy

First Nations Media Australia:

... the Productivity Commission has an opportunity to develop collaborative evaluation systems and practices that benefit both Government and Aboriginal and Torres Strait Islander people through evaluating programs in partnership with First Nations organisations and communities, so that Aboriginal and Torres Strait Islander people can use information gathered about them for their own purposes. (sub. 30, p. 11)

Victorian Aboriginal Child Care Agency:

Consistency in approaches is important ... Additional guidance on how the principles should be enacted will be needed to ensure that agencies abide by the principles in commissioning and implementing evaluations. ... There is an opportunity for the Productivity Commission to recognise and support the need for further development of recognition for Aboriginal ways of knowing and doing in relation to policy and program evaluation. (sub. 26, pp. 2–4)

Queensland Nurses and Midwives' Union:

We see the development of an evaluation strategy for Aboriginal and Torres Strait Islander people as an opportunity to embed transparency into programs and policies. This will better inform the beneficiaries of the policy and ensure the programs and policies are working and achieving what they set out to do. (sub. 12, p. 3)

The National Mental Health Commission:

... the development of a targeted Indigenous Evaluation Strategy has the potential to significantly improve the quality and outcomes of programs that serve Aboriginal and Torres Strait Islander peoples. (sub. 19, p. 1)

Fiona Allison and LawRight:

There are currently insufficient Indigenous evaluations being conducted. We ... support development of an overarching Indigenous Evaluation Strategy that will raise awareness of, encourage and guide good practice in this area — though the utility of this strategy will be wholly dependent on the way in which it is implemented. (sub. 18, p. 3)

Royal Australasian College of Physicians:

We therefore welcome the development of a long overdue whole-of-government evaluation strategy for Indigenous policy. This will be critical to effective and cost-effective program design, delivery, and accountability. (sub. 77, p. 2)

The Aboriginal Health Council of Western Australia:

... the evaluation framework provides a unique opportunity to shift the dialogue on Aboriginal health from a deficit to a strengths-based model. Too often, data demonstrating Aboriginal health and wellbeing focus on deficits and limitations when, instead, data should be reframed to reflect achievements in health outcomes. (sub. 42, p. 4)

1.4 How we approached the task

The ultimate objective of any policy — in this case a whole-of-government Indigenous Evaluation Strategy — is to improve the wellbeing of the community overall. Because all policies have costs, to be convinced that a policy will ‘make a difference’ it is necessary to demonstrate that the benefits of the policy outweigh the costs to the Australian community. The Commission’s approach to developing the Strategy therefore involved looking at the

benefits and impacts of the current arrangements for evaluation, and alternative approaches, to the lives of Aboriginal and Torres Strait Islander people, and Australians more generally, in light of the costs — that is, using a wellbeing approach.

When thinking about the benefits and costs of a Strategy, the Commission reflected on a number of issues, including:

- the current state of play for evaluation across the Australian Public Service: what are the strengths and weaknesses of the current arrangements? To what extent are evaluations being undertaken? Are evaluation findings influencing policy decisions and practice and providing insights for priority areas? If not, why not? How is the accumulated evaluation knowledge used? Could it be better used?
- the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people: what is the current quality of evaluations and can it be improved on? Are the right policy and program areas being prioritised for evaluation? Are the priorities of Aboriginal and Torres Strait Islander people reflected in determining evaluation priorities? What do we know about what works or does not work? To what extent have past evaluations delivered benefits to Aboriginal and Torres Strait Islander people?
- what does quality evaluation look like? Can evaluation practices be improved, including the feedback loops within and across government for using evaluation findings in policy and program design, and the way results of evaluations and subsequent actions are communicated to evaluation users and Aboriginal and Torres Strait Islander people? Is there a role for greater oversight of evaluations?
- what data, evaluation capabilities and resources are required to produce and use high quality evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people?

To help answer some of these questions, the Commission released an issues paper in June 2019. We received 112 submissions (including three conducted orally) and three brief comments in response to the Issues Paper (figure 1.1). We also sent a questionnaire to all Australian Government agencies asking for information on evaluation policies and practices and for any evaluations undertaken by agencies in the previous three years (for more information on the questionnaire see chapter 3 and appendix B).

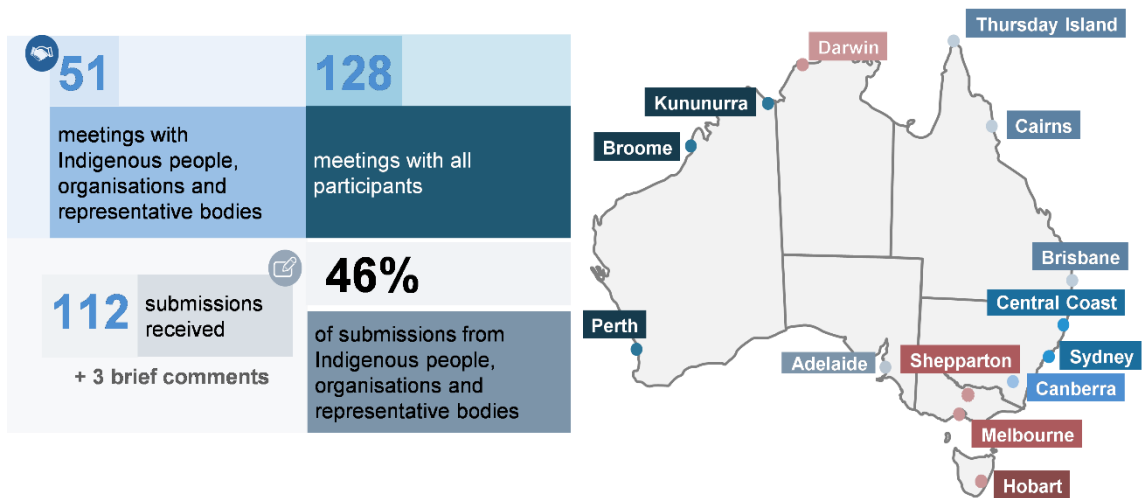
The draft Strategy is also informed by the literature on good practice evaluation approaches and methods, ethics and research guidelines, standards and strategies, and governance arrangements for evaluation adopted in other similar countries (New Zealand, Canada, the United States and the United Kingdom).

Some participants⁵ suggested that Aboriginal and Torres Strait Islander people should be co-designing the Strategy. There are, however, some constraints around co-designing the

⁵ For example, Generation One, Minderoo Foundation (sub. 5, p. 1), Aboriginal Affairs NSW (sub. 70, p. 2), National Aboriginal Community Controlled Health Organisation (sub. 95, p. 2).

Strategy as the Australian Government has asked the Productivity Commission to develop the Strategy. That said, we are consulting widely in developing the Strategy, including with Aboriginal and Torres Strait Islander people, communities and organisations (figure 1.1). The Commission is also cognisant of the changing policy landscape, including the formation of the Joint Council on Closing the Gap and the development of priority reforms to be included in a new National Agreement on Closing the Gap (box 1.4). These co-design processes informed and influenced the development of the draft Strategy.

Figure 1.1 The Commission engaged widely across Australia



The Commission also engaged with Australian Government and state and territory government agencies (including participating in a forum hosted by Aboriginal Affairs NSW on what they had learnt about evaluation practice) while developing the draft Strategy. The Commission also held two roundtables on the topics of:

- An Indigenous Evaluation Strategy: Objectives, Principles and Defining Success
- Culture, Capability and Governance in Government for Better Outcomes Through Evaluation.

The roundtables were used to test the Commission's thinking on aspects of the Strategy with Aboriginal and Torres Strait Islander organisations, academics and government agencies.

Appendix A includes a list of project participants and submissions.

The Strategy is a draft and the Commission is seeking views on the objective, the principles and practical guidance and the proposed governance arrangements for the Strategy from Aboriginal and Torres Strait Islander people and organisations, government agencies, academics, evaluation practitioners, policy or governance experts and others who will be affected by the Strategy.

The Strategy will be a ‘living’ document subject to ongoing review (chapter 10). It will be critical to the long-term success of the Strategy that Aboriginal and Torres Strait Islander people and Australian Government agencies are involved in the review process.

What policies and programs are covered by the Strategy?

The Strategy will cover all policies and programs that affect Aboriginal and Torres Strait Islander people (that is, both ‘mainstream’ programs — universal programs provided for all Australians) and Indigenous-specific programs (programs provided for Aboriginal and Torres Strait Islander people specifically). Given four in every five dollars of direct government expenditure on Aboriginal and Torres Strait Islander people is spent on mainstream programs (box 1.6), including mainstream programs in the Strategy is imperative and was strongly supported by participants (box 1.7).

Box 1.6 Expenditure on Aboriginal and Torres Strait Islander people

Government expenditure on mainstream services accounted for 82 per cent of direct government expenditure on Aboriginal and Torres Strait Islander people in 2015-16, with Indigenous-specific services accounting for the remaining 18 per cent.

In 2015-16, the most recent year for which data on this expenditure breakdown are available (from the 2017 *Indigenous Expenditure Report*), total direct government expenditure^a on Aboriginal and Torres Strait Islander people was estimated to be \$33.4 billion, with the estimated direct expenditure per person being \$44 886 for Aboriginal and Torres Strait Islander people, about twice the rate for non-Indigenous Australians (\$22 356). The higher per person expenditure for Aboriginal and Torres Strait Islander people reflects the combined effects of:

- greater *intensity of service use* by Aboriginal and Torres Strait Islander people (representing \$14 349 or 63.7 per cent) — reflecting the greater needs of and the younger age profile of the Aboriginal and Torres Strait Islander population
- higher *cost of providing services* (representing \$8181 or 36.3 per cent) — for example, because of more remote locations, or because of provision of targeted services in addition to mainstream services (for example, Indigenous liaison officers in hospitals) to ensure provision of appropriate and accessible services.

The Australian Government accounted for \$14.7 billion (44 per cent) of direct government expenditure on Aboriginal and Torres Strait Islander people in 2015-16, with \$18.8 billion (56 per cent) being provided by state and territory governments.

^a Direct expenditure refers to government expenditure on services and payments provided directly to individuals, non-government service providers and local governments. Transfers between governments (indirect expenditure) are excluded to avoid double counting.

Source: SCRGSP (2017).

The Indigenous Evaluation Strategy will only apply to Australian Government agencies — the Australian Government accounts for about 44 per cent of direct government expenditure on Aboriginal and Torres Strait Islander people. However, others, such as Aboriginal and Torres Strait Islander organisations, state, territory and local governments, and not-for-profit

entities will also be affected as stakeholders in policy implementation, delivery and evaluation (and we have drawn heavily on their knowledge and experience).

Box 1.7 Participants pointed to the importance of including mainstream policies and programs under the Strategy

The National Mental Health Commission:

Aboriginal and Torres Strait Islander peoples' experiences with mainstream services are likely to differ significantly from the broader population, due to disparities in health and wellbeing status and specific barriers such as racism and lack of cultural safety. Ensuring that these unique experiences are captured in evaluations of mainstream services could help to facilitate improvements in experiences and outcomes for Aboriginal and Torres Strait Islander peoples. (sub. 19, p. 1)

Tony Dreise et al:

... the Commission recognises and acknowledges the impact of mainstream programs on First Nations people (whether those programs and policies are explicitly directed to First Nations people or not). These programs impact First Nations people in fundamental ways ... and thus should be a major focus for the Commission in developing the IES. While we encourage the Commission to consider including evaluations of mainstream policies and programs within the scope of the IES, we suggest that any such evaluations focus on how mainstream policies and programs affect First Nations peoples specifically. (sub. 33, p. 1)

Queensland Aboriginal and Islander Health Council:

Mainstream programs should require evaluation and accountability to Aboriginal and Torres Strait Islander representative bodies. The Productivity Commission correctly identifies that the vast majority of Australian Government funding towards Aboriginal and Torres Strait Islander services is spent through mainstream programs and services. ... Yet, these services are often the least evaluated. (sub. 55, p. 7)

The Western Australian Government:

The Strategy should not be limited to Indigenous-specific Australian Government policies and programs. Too often, the impact of mainstream policies and programs on Aboriginal and Torres Strait Islander people is not adequately considered in evaluation. This is particularly of concern, as 'four in every five dollars spent by the Australian Government in providing services to Aboriginal and Torres Strait Islander people is spent through mainstream programs and services ...' (sub. 74, p. 2)

1.5 Improving lives is the objective

There is agreement that the objective of the Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people (box 1.8). With this as the objective of the Strategy, it is important to understand what Aboriginal and Torres Strait Islander people value and what is important to their lives.

Some of the common themes from discussions, submissions and the literature about what is important for Aboriginal and Torres Strait Islander wellbeing (box 1.9) are relationships, connection to Country, community and culture (including a sense of belonging), empowerment and self-determination (the right to make decisions on matters that affect the lives and communities of Aboriginal and Torres Strait Islander people).

Box 1.8 **An objective to improve the lives or outcomes of Aboriginal and Torres Strait Islander people**

Empowered Communities:

Improving outcomes for Indigenous people is the right overarching objective for the Indigenous Evaluation Strategy. (sub. 41, p. 6)

Australian Institute of Aboriginal and Torres Strait Islander Studies:

The clear objective of all government action that impacts Aboriginal and Torres Strait Islander people should be to improve wellbeing, to ensure that Indigenous people have the capabilities and opportunities to live the life they value, in a society that values and affirms Indigenous peoples' identities, cultures and contributions to Australian nationhood. (sub. 72, p. 3)

The National Indigenous Australians Agency:

... the framework that underpins this strategy should be broad enough to be applied to mainstream government programs and have at its core the objective to improve the lives of Indigenous Australians. (sub. 86, p. 8)

The Tasmanian Government:

The Tasmanian Government suggests that the Indigenous Evaluation Strategy aim to ensure that evaluation of Aboriginal and Torres Strait Islander-specific policies and programs adopts a 'do no harm' approach to individual and community health and wellbeing by avoiding practices that may entrench negative stereotypes and deficit models. The purposes of evaluation could include advancing the redress of Aboriginal and Torres Strait Islander inequality, and the continued enhancement of the inherent rights, cultures and traditions of Aboriginal and Torres Strait Islander people. (sub. 100, p. 1)

Some participants said historical events (box 1.10)⁶ (including the impact of European settlement and past government action on Aboriginal and Torres Strait Islander people's connection to Country and culture) need to be acknowledged when undertaking evaluations because of their effect on lives and communities. For example, the Victorian Aboriginal Community Controlled Health Organisation said:

... racism, discrimination and trauma live on in Aboriginal Communities, causing real impacts on health and wellbeing. There must be an ongoing acknowledgement by the Government at all levels in order to foster genuine evaluations of the policies and programs affecting Aboriginal people and Communities. (sub. 44, p. 8)

⁶ A more detailed discussion on the history of Aboriginal and Torres Strait Islander people since European settlement is contained in chapter 1 of the Commission's *Overcoming Indigenous Disadvantage Report, Key Indicators 2016*.

Box 1.9 **Some insights on what Aboriginal and Torres Strait Islander people value and what is important for their wellbeing**

Aurora Milroy & the ANZSOG First Peoples Team, in a submission to the Australian Public Service Review, pointed to some key Indigenous values and concepts:

Caring for Country: The land is our mother and provides for us. In turn, we have an obligation to care for land, maintain ceremony, and ensure the long-term wellbeing of Country.

Kinship: As with Country, we also have responsibilities to care for our broader kinship network, ensure balance and act reciprocally.

Consensus decision-making: The interconnection of all things requires us to work collaboratively to make decisions.

Process is important: We value making decisions in the right way, including by involving all relevant parties in the process from the start. The right outcome is one that is achieved through the right means.

Honesty: The long-term sustainability of our systems and relationships relies on trust and honesty. If you see everything as connected, the failure to be honest in your interactions with others threatens the strength of our relationships, and therefore the whole system.

Humility: We do not have all the answers and must recognise when others have more authority or information to help us achieve the best outcome. (2019, p. 7)

Yap and Yu similarly defined Indigenous well-being as:

Autonomy and self-determination for Indigenous peoples ... A second theme in the literature points to Indigenous wellbeing being strongly tied to Indigenous peoples' connection to country and culture. ... The central importance of family and kinship for the wellbeing of Indigenous people cannot be over-emphasised. For most Indigenous people, their wellbeing is linked to their community, country and sense of belonging. (Yap and Yu 2016b, pp. 317–321)

The National Health and Medical Research Council (2018a, p. 7) noted that Aboriginal and Torres Strait Islander People and communities each have their own established and respected values and protocols, and unique ways of expressing their different values. It identified six core values as being important to all Aboriginal and Torres Strait Islander People — reciprocity, respect, equity, cultural continuity, spirit and integrity, and responsibility.

A survey of 842 Aboriginal people, aged 15-34 years from four different Aboriginal communities in remote Australia, found that culture, empowerment and community play key roles in the interplay with education, employment and health, as part of a holistic and quantifiable system of wellbeing (Cairney et al. 2017).

The Yawuru Wellbeing Project highlighted the importance of relationships to good *liyan* (for Yawuru people, *mabu liyan* is at the heart of what it is to have and know a good life). Connectedness to family, community and country were all recurrent themes. Good health and having a basic standard of living were also identified as substantial elements of *liyan* (where health included physical health as well as social, emotional and spiritual wellbeing) as was self-determination (having a say on things affect families, community and country and feeling respected and being free from discrimination) (Yap and Yu 2016a).

Box 1.10 **Historical factors provide important context**

Aboriginal people have lived in Australia for 60 000 years, and Torres Strait Islander people settled some 10 000 years ago. Prior to European settlement, it is estimated that there were around 250 distinct nations, with different languages and social systems. Much of this diversity remains today.

Attachment to the land was, and still is, a central element of Aboriginal and Torres Strait Islander people's cultures, customs and laws. However, when the British established the penal colony of New South Wales in January 1788, the legal system operated as if Australia belonged to no one, and denied that Aboriginal and Torres Strait Islander people had sovereignty or property rights over the land. Aboriginal and Torres Strait Islander people were dispossessed of their traditional lands, largely through violence and murder.

From 1860, 'protective' legislation (known as the 'Aborigines Acts') required many Aboriginal people to live on reserves — with lives on the reserves being highly regulated and basic human rights suppressed. Aboriginal people living outside reserves also had their rights restricted.

In the 1900s, governments adopted a more assimilationist approach, with explicit goals of integrating Aboriginal and Torres Strait Islander people into 'mainstream' society. Between 10 and 30 per cent of Aboriginal and Torres Strait Islander children were forcibly removed from their families and communities between 1910 and 1970, now commonly referred to as 'the Stolen Generations'.

The 1967 referendum, allowing the Australian Government to legislate with respect to Aboriginal people, and to include Aboriginal people in the Census, was a watershed in relations between Aboriginal and Torres Strait Islander people and non-Indigenous Australians, and was supported by over 90 per cent of voters (the largest majority for any Australian referendum).

The *Racial Discrimination Act 1975* (Cwlth) was introduced in 1975, making racial discrimination unlawful in certain contexts, including employment, housing, access to goods and services, and access to public places.

In 1990, an elected Aboriginal and Torres Strait Islander Commission (ATSIC) was established to provide advice to the government and to deliver some services (although not mainstream services, and its level of responsibility for Indigenous-specific services declined over time). ATSIC was abolished in 2004, and responsibility for services delivered by ATSIC returned to line departments.

Two major national inquiries helped promote a deeper understanding of the issues facing Aboriginal and Torres Strait Islander people. The Royal Commission into Aboriginal Deaths in Custody (1987 – 1991) reported on the complex effects of dispossession, colonisation and institutional racism on Aboriginal peoples, while a national inquiry into the separation of Indigenous children from their families (1995 – 1996) resulted in the 1997 *Bringing Them Home* report on the Stolen Generations. In 2008, the Prime Minister of Australia, with bipartisan support, made a formal apology to the Stolen Generations.

Source: SCRGSP (2016).

Wellbeing frameworks for Indigenous people

Aboriginal and Torres Strait Islander people should define what a good life, or success, looks like for them. As Yap and Yu said:

Whilst achieving Indigenous wellbeing is an important outcome in its own right, a critical factor often overlooked is that Indigenous participation and decision making in the process of determining what wellbeing means for them is an essential and fundamental aspect of Indigenous wellbeing. This goes to the heart of self-determination over one's life and respects the principles outlined in the United Nations Declaration on the Rights of Indigenous Peoples. (Yap and Yu 2016a, p. 98)

Two examples of wellbeing frameworks led by Aboriginal and Torres Strait Islander people and representatives are:

- The Dance of Life framework — this framework reflects a multi-dimensional model of health and wellbeing from an Aboriginal and Torres Strait Islander perspective. Represented by paintings by Helen Milroy, the framework consists of five dimensions — cultural, spiritual, social, emotional and physical — within which are a number of layers that reflect historical, traditional and contemporary influences. The concept emphasises the intersection of the layers and dimensions that create the interconnectedness for a whole-of-life approach to Aboriginal and Torres Strait Islander health and wellbeing (RANZCP 2020).
- The Gayaa Dhuwi (Proud Spirit) Declaration, which was launched by National Aboriginal and Torres Strait Islander Leadership in Mental Health (NATSILMH) in 2015. It is designed as a companion declaration to the internationally recognised Wharerātā Declaration, and has five key elements, underpinned by nine principles of Aboriginal and Torres Strait Islander social and emotional wellbeing (box 1.11) (NATSILMH 2015).

The Australian Bureau of Statistics (ABS) has also developed an Indigenous wellbeing framework in conjunction with Aboriginal and Torres Strait Islander people. The framework has nine domains (figure 1.2) which together attempt to provide a holistic view of Indigenous wellbeing (ABS 2010).

A diverse population

There is considerable linguistic and cultural diversity amongst Aboriginal and Torres Strait Islander people, with communities each having established values and protocols, and unique way of expressing their values. This diversity also extends to histories and experiences of Aboriginal and Torres Strait Islander people and communities (NHMRC 2018a, p. 1). As Walter and Andersen said:

... the diversity of our histories, our contemporary and historical experiences of colonization, our traditional and contemporary cultural lives, and our current social, political, and economic positioning mean that often we are more different than we are similar. (2013, p. 62)

Box 1.11 Principles of the Gayaa Dhuwi Declaration

1. Aboriginal and Torres Strait Islander health is viewed in a holistic context, that encompasses mental health and physical, cultural and spiritual health. Land and sea is central to wellbeing. When the harmony of these interrelations is disrupted, Aboriginal and Torres Strait Islander ill-health will persist.
2. Self-determination is central to the provision of Aboriginal and Torres Strait Islander health services.
3. Culturally valid understandings must shape the provision of services and must guide assessment, care and management of Aboriginal and Torres Strait Islander peoples' health problems generally, and mental health problems, in particular.
4. It must be recognised that the experiences of trauma and loss, present since European invasion, are a direct outcome of the disruption to cultural wellbeing. Trauma and loss of this magnitude continues to have inter-generational effects.
5. The human rights of Aboriginal and Torres Strait Islander peoples must be recognised and respected. Failure to respect these human rights constitutes continuous disruption to mental health (versus mental ill-health). Human rights relevant to mental illness must be specifically addressed.
6. Racism, stigma, environmental adversity and social disadvantage constitute ongoing stressors and have negative impacts on Aboriginal and Torres Strait Islander peoples' mental health and wellbeing.
7. The centrality of Aboriginal and Torres Strait Islander family and kinship must be recognised as well as the broader concepts of family and the bonds of reciprocal affection, responsibility and sharing.
8. There is no single Aboriginal or Torres Strait Islander culture or group, but numerous groupings, languages, kinships, and tribes, as well as ways of living. Furthermore, Aboriginal and Torres Strait Islander peoples may currently live in urban, rural or remote settings, in traditional or other lifestyles, and frequently move between these ways of living.
9. It must be recognised that Aboriginal and Torres Strait Islander peoples have great strengths, creativity and endurance and a deep understanding of the relationships between human beings and their environment.

Source: NATSILMH (2015).

Figure 1.2 **ABS wellbeing framework for Aboriginal and Torres Strait Islander people**



Source: ABS (2010).

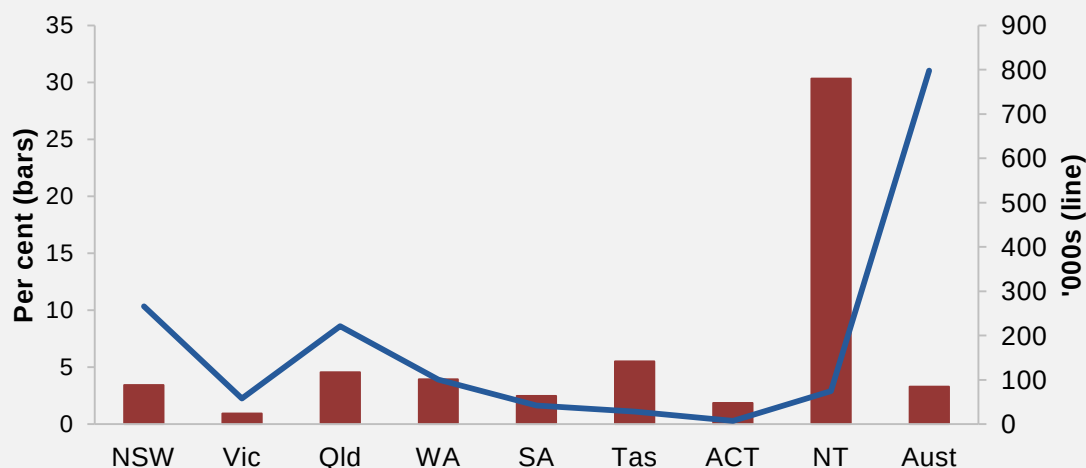
The Aboriginal and Torres Strait Islander population is also geographically diverse (box 1.12). While the conversation about evaluating policies and programs affecting Aboriginal and Torres Strait Islander people is often focused on those living in remote communities, most of the Aboriginal and Torres Strait Islander population (more than 80 per cent) live in major urban or regional cities. Maggie Walter made this very point:

The focus of the government's Indigenous Advancement Strategy and Indigenous policy is very much on remote people, which is understandable, given some of the issues of poor housing, health, education and outcomes. But this comes at the exclusion of the 80 per cent of Aboriginal and Torres Strait Islander people who do not live in remote areas. And this is evident in the data available. (sub. 112, p. 6)

Outcomes and opportunities for Aboriginal and Torres Strait Islander people can also be very different between urban, regional and remote locations. For example, the circumstances of Aboriginal and Torres Strait Islander people living in inner Sydney are likely to be very different to those living in Central Australia.

Box 1.12 Key Aboriginal and Torres Strait Islander demographics

Aboriginal and Torres Strait Islander people make up about 3.3 per cent of the total Australian population. The proportion of Aboriginal and Torres Strait Islander people in each state and territory varies from less than 1 per cent of the population in the Australian Capital Territory and Victoria to 30 per cent in the Northern Territory. The figure below shows the number of Aboriginal and Torres Strait Islander people by state and territory, as at 30 June 2016.



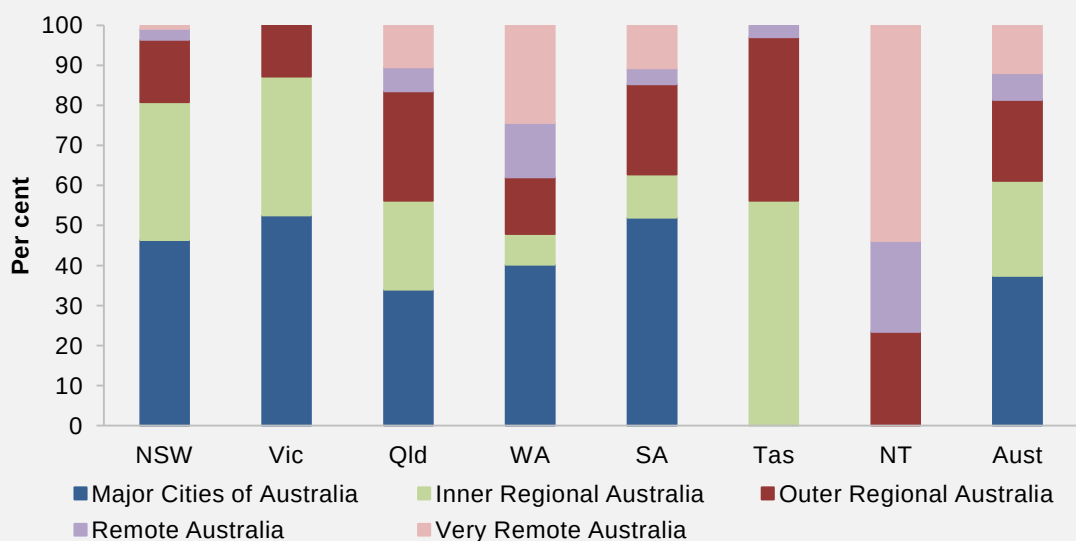
Projections data for the Aboriginal and Torres Strait Islander population and Estimated Resident Population for 2016 (used to calculate proportions) are not comparable as they are sourced from different data collections.

Source: SCRGSP (2019).

Most of the Aboriginal and Torres Strait Islander population live in major urban or regional cities (81.3 per cent). While 18.7 per cent of the Aboriginal and Torres Strait Islander population live in remote areas, this proportion is 76.6 per cent in the Northern Territory and 38 per cent in Western Australia. The figure below shows where Aboriginal and Torres Strait Islander people live, as at 30 June 2016 (although data by remoteness area was not available for the ACT).

(continued next page)

Box 1.12 (continued)



Data source: ABS Estimates of Aboriginal and Torres Strait Islander Australians, June 2016 (cat. no. 3238.0.55.001).

The Central Australian Aboriginal Congress highlighted the need for ‘evaluation of policies or programs affecting Aboriginal people to include the views and experiences of Aboriginal community members, leaders, and organisations and reflect the diversity of Aboriginal Australia with respect to gender, remoteness, culture, age, disability, sexuality’ (sub. 48, p. 6). The Smith Family also pointed to the importance of the Strategy acknowledging ‘the diversity of characteristics, experiences and perspectives of Aboriginal and Torres Strait Islander peoples’ and the ‘different experiences, by for example, socioeconomic background, location and gender’ (sub. 56, p. 5).

In sum, there is a lot of diversity among Aboriginal and Torres Strait Islander people and communities. The Strategy needs to acknowledge this diversity of views, perspectives and experiences.

Strategy success involves balancing the interests of all those affected

While there is widespread agreement that the objective of the Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people, by improving the quality and use of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, the Strategy also needs to recognise and acknowledge that other people and organisations will be affected (figure 1.3) and that their interests need to be considered.

Figure 1.3 Who will be affected by the Strategy?



In the context of what Strategy success would look like, a common theme was the importance of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in program and policy evaluation planning, design, implementation and priority setting (box 1.13).

Many participants called for the Articles in the *United Nations Declaration on the Rights of Indigenous Peoples* (UNDRIP) (box 1.14), which was endorsed by the Australian Government in 2009, to be recognised in the Strategy. The declaration outlines minimum standards for the ‘survival, dignity and wellbeing of the indigenous people of the world’ and the right to self-determination.

While self-determination means different things to different people, the National Health and Medical and Research Council explained the right to self-determination to mean that Aboriginal and Torres Strait Islander Peoples ‘have the freedom to live well and to live according to their values and beliefs’ (NHMRC 2020). There are four main ways in which Aboriginal and Torres Strait Islander people exercise self-determination:

- having a choice in determining their development paths and how their lives are governed
- participating in decisions that affect their lives
- having control over their lives and future including economic, social and cultural development
- having the means by which every person might achieve their full potential (NHMRC 2020).

In the context of evaluation this can mean people being able to make their own decisions about whether to participate in an evaluation and Aboriginal and Torres Strait Islander people working in partnership with evaluators to ensure Indigenous values and knowledges are central to the evaluation.

Box 1.13 Participants highlighted the importance of incorporating Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges

Australian Evaluation Society:

A community partnership approach would enable Indigenous community representatives to oversight and advise on evaluation approaches and practice (i.e. both overall and on a case by case basis). Engaging Indigenous people in co-design processes, would encourage self-determination and valuing of Aboriginal and Torres Strait Islander knowledges. (sub. 49, p. 18)

The Smith Family:

An Indigenous Evaluation Strategy must rest on the principle of self-determination ... This principle implies the importance of Aboriginal and Torres Strait Islander people fully participating in the development, implementation and management of programs and policies that affect them, including the development of an Indigenous Evaluation Strategy and evaluation activities which flow from it. (sub. 56, p. 3)

Australian Institute of Aboriginal and Torres Strait Islander Studies:

Any strategy evaluating policies and programs affecting Aboriginal and Torres Strait Islander peoples should measure success by the extent to which they improve or promote the enjoyment of rights and Australia's responsibilities under the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). ... The principle of self determination requires that Indigenous peoples be involved in decisions that affect them, including the design, delivery and evaluation of government policies and programs. (sub. 72, p. 3)

National Family Violence Prevention Legal Services Forum:

For evaluation to genuinely benefit Aboriginal and Torres Strait Islander communities there needs to be an intention to serve an agenda which is determined by Aboriginal and Torres Strait Islander people. (sub. 66, p. 7)

Reconciliation Australia:

It is the recommendation of Reconciliation Australia, that a co-design process for the development of the Indigenous Evaluation Strategy is employed to ensure regular and structured engagement with Aboriginal and Torres Strait Islander peoples and peak bodies. (sub. 79, p. 8)

The Western Australian Government:

...the strategy should aim to maximise Aboriginal engagement in evaluation design and implementation, by focusing on genuine partnership and co-design ... (sub. 74, p. 1)

Another theme in the context of Strategy success is the importance of better using evaluation findings to learn and drive improvements, and to inform future government spending (providing confidence that we are getting the biggest 'bang for buck'). Many also spoke about the importance of knowledge management, making evidence accessible and the sharing of evaluation data, results and lessons (with a particular focus on translation for Aboriginal and Torres Strait Islander communities and policy makers).

Box 1.14 United Nations Declaration on the Rights of Indigenous Peoples and self-determination

Aboriginal and Torres Strait Islander people have for decades fought for self-determination and their right to make decisions on matters that affect their lives and communities.

The *United Nations Declaration on the Rights of Indigenous Peoples* (UNDRIP), adopted on 13 September 2007, with Australia endorsing it on 3 April 2009, establishes a universal framework of minimum standards for the survival, dignity, security and wellbeing of indigenous peoples. While not legally binding, the Declaration is significant because indigenous peoples, including Aboriginal and Torres Strait Islander peoples, were involved in its drafting.

The Declaration has six foundational rights: individual and collective access to human rights; equality and freedom from discrimination; self-determination; self-governance; participation in the life of the State; and nationality. It also outlines rights to land and resources, education and information, cultural and spiritual identity and indigenous-owned institutions.

Beyond the foundational rights, there are several rights outlined in the Declaration that are of particular relevance to the Indigenous Evaluation Strategy. These include the right of indigenous peoples to:

- participate in decision-making affecting indigenous peoples (article 18)
- determine and develop priorities and strategies to exercise their right to development and be actively involved in developing, determining and administering programs (article 23)
- promote, develop and maintain institutional structures (article 34) for their own development and decision making (articles 18 and 20).

While self-determination can mean different things to different people, and there is no predetermined outcome of what self-determination looks like, the Declaration describes self-determination as the ability for Indigenous people to ‘freely determine their political status and freely pursue their economic, social and cultural development’. And in exercising this right they have the right to ‘autonomy or self-government’. The right to self-determination is seen as a ‘foundational right’: the other rights cannot be fully enjoyed without it, while furthering the other UNDRIP rights also furthers self-determination.

Sources: AHRC (2010); UN (2007).

The main feedback from government agencies was that they wanted the Strategy to provide practical guidance on evaluation approaches and methodologies, including advice on how to ethically engage with, and respect the perspectives of, Aboriginal and Torres Strait Islander people and communities. For example, the Department of Health said:

The framework could usefully include advice about rigour and participatory style approaches in the design of evaluations. It would be helpful for the Strategy to offer guidance to Australian Government agencies about how to engage well and ethically with Indigenous stakeholders and communities. The evaluation co-design space is a new area in evaluation, and guidance about how and when to apply co-design to evaluations would also be useful. (sub. 35, p. 4)

And the Tasmanian Government said it considered:

... there may be value in including a component to guide agencies on what activities and structures they can put in place to enable effective evaluations. ... by not only identifying the

processes and institutional characteristics needed to promote the adoption and success of the Indigenous Evaluation Strategy but also encouraging and helping agencies to implement them. (sub. 100, p. 2)

For the Strategy to achieve its objective of improving the lives of Aboriginal and Torres Strait Islander people, it needs to be more than a document that sets out objectives, principles and priorities for selecting and conducting evaluations. The Strategy needs to support the government working in a different way with Aboriginal and Torres Strait Islander people, put in place governance arrangements that centre Aboriginal and Torres Strait Islander people, and embed incentives to learn from and use evaluation findings in policy decision making and program improvements. And it needs to be more than a compliance exercise or a box to be ticked by agencies. It needs to be about improving policy through conducting high quality evaluations and then using evaluation findings in policy decision making. It is evaluation's contribution to better policies, not the undertaking of evaluation per se, that will improve the lives of Aboriginal and Torres Strait Islander people.

1.6 A roadmap for the rest of the report

The next chapter looks at past and present evaluation practice across the Australian Public Service.

Chapter 3 discusses the current state of the evaluation evidence base, including what policies and programs affecting Aboriginal and Torres Strait Islander people have been evaluated, the quality of those evaluations, what we have learned from evaluation, and the gaps in our knowledge that could be filled by better evaluation.

Chapter 4 discusses the various types, approaches of methods of evaluation available to agencies.

Chapter 5 looks at the practice of evaluation, focusing on how to achieve high-quality evaluation.

Chapter 6 looks at evaluation priorities, and how they should be determined.

Chapter 7 discusses governance arrangements for encouraging a culture of learning and evaluation within agencies.

Chapter 8 looks at how the capability to produce and use evaluations might be strengthened.

Chapter 9 looks at how data are used in evaluation, the issues surrounding its use and how the use of data could be improved.

Chapter 10 discusses arrangements for monitoring and reviewing the Strategy, including how to assess the performance of agencies against the Strategy and whether the Strategy remains fit-for-purpose and capable of delivering on its overall objective.

Chapter 11 brings together insights from earlier chapters and outlines the key components and proposed principles for the Strategy.

2 Evaluation in the APS — history and current practice

Key points

- Evaluation is currently decentralised within the Australian Public Service (APS). Evaluation in the APS has had a varied history — over time evaluation arrangements have been formalised and centrally supported, then wound back, with agencies given autonomy around how they plan, undertake and support evaluation.
- There are currently no formal requirements for Australian Government agencies to evaluate policies and programs. There is also no:
 - government-wide evaluation strategy or policy to guide evaluation across the APS — either generally or specifically for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people
 - process to set evaluation priorities at a whole-of-government level
 - requirement to publish evaluations or a central register or repository where evaluation findings are published.
- The lack of a formal, systematic and whole-of-government approach to evaluation means that the way and extent that Australian Government agencies plan for, undertake, fund, use and share evaluations varies.

This chapter looks at the history and current state of evaluation across the Australian Public Service (APS). It is important to understand the environment in which Australian Government agencies will be operating in as they seek to comply with the Indigenous Evaluation Strategy (the Strategy). And, in the context of what governance arrangements might be required to support the Strategy, it is important to know what has (or has not) worked in the past to support evaluation activity. The chapter also looks at how governments approach evaluating policies and programs affecting Aboriginal and Torres Strait Islander people.

2.1 The ebbs and flow of evaluation history

Evaluation of policies and programs in the APS has a varied history, with evaluation arrangements formalised and centrally supported, before being wound back, with Australian Government agencies given autonomy around how they plan, undertake and support evaluation. Some recent proposed changes to how evaluation is done in the APS point to a move back to a more centralised approach.

A brief history

Efforts by Australian governments to promote evaluation can be traced back to at least the 1950s and 1960s (Sharp 2003), however, it was two reports in the 1970s — the Coombs⁷ and Baume reports — that argued for a greater focus on evaluation in the APS.

The Coombs report recommended a central evaluation unit (within the Department of the Prime Minister and Cabinet) be established to ‘evaluate the effectiveness of ongoing programs against the general philosophy, policies and priorities of the government’ (1976, p. 385). It also paid particular attention to the need for ‘review and assessment’ of government programs in the policy area of ‘Aboriginal affairs’:

This Report has emphasised frequently the need for the progressive assessment of government programs. Nowhere is it more acute than in Aboriginal affairs. The complexity of the problems and the unexpectedness of many of the difficulties which have been encountered in carrying out apparently well-founded programs emphasise this need. (RCAGA 1976, p. 342)

The Senate Standing Committee on Social Welfare report, *Through a Glass, Darkly* (also known as the Baume Report), looked at the role of evaluation in human services and concluded that:

Evaluation is one tool to help us understand what is going on. It is a tool becoming better appreciated and recognised in Australia today, but it is still far too little applied and seldom funded. Further, there is not an appropriate context — nor are there the structures and processes needed — for development of evaluation activity. (SSCSW 1979, p. 1)

Reforms to the APS in the 1980s saw a greater focus placed on program management and budgeting as well as monitoring and evaluation. The Department of Finance (DoF) was given a central role to support improvements in evaluation capabilities and, in 1986, published an evaluation handbook. In 1988, the Government required departments to submit monitoring and evaluation proposals to the DoF, and all New Policy Proposals (NPPs) had to include performance measures and an evaluation plan (Mackay 2011b, p. 5).

The late 1980s also saw a change to the arrangements for the evaluation of programs and policies affecting Aboriginal and Torres Strait Islander people — responsibility for evaluating policies and programs targeted at Aboriginal and Torres Strait Islander people moved from the (then) Department of Aboriginal Affairs (where it had been housed since 1979) to a new, quasi-independent body, the Office of Evaluation and Audit (box 2.1).

⁷ The Coombs Royal Commission on Australian Government Administration.

Box 2.1 Office of Evaluation and Audit

The Office of Evaluation and Audit (OEA) was established under the *Aboriginal and Torres Strait Islander Commission Act 1989* (Cwlth) and was tasked with the regular evaluation and audit of Aboriginal and Torres Strait Islander Commission (ATSIC) operations and delivering on special ministerial requests.

The Act set out how the OEA should approach its evaluation functions, calling for regular evaluations on all programs conducted and funded under the Act, as well as overseeing the operations of every ATSIC office. Evaluations were to be delivered at least once every three years and operations came to reflect the evaluation environment developing under the Commonwealth Evaluation Strategy.

The Director of the OEA was appointed by the responsible minister, maintaining independence from the ATSIC Board of Commissioners, and the OEA reported directly to Parliament.

When ATSIC was abolished in 2005, the operations of the OEA moved to the Department of Finance before being incorporated into the Australian National Audit Office in 2009.

Evaluation activity within the OEA from its inception remained fairly consistent, but output of completed program evaluation reports in a year never exceeded eight. While located within ATSIC and the Department of Finance all evaluation activity was publically reported in the Annual Report of its parent-agency.

Sources: *Aboriginal and Torres Strait Islander Commission Act 1989* (Cwlth); Commission estimates based on ATSIC Annual Reports (1992–2005) and Department of Finance Annual Reports (2005–2009); Dillon (1992).

A 1988 review of evaluation activities across the public service under the Financial Management Improvement Program (FMIP) found evaluation processes not integrated into financial decision making of agencies, and the focus of performance information remained largely on efficiency and process rather than effectiveness. It also found that the conduct of evaluation was hampered by a poor level of evaluation skill and analytical capacity in agencies and that the role the DoF played encouraging or enforcing evaluation practice remained unclear (House of Representatives Standing Committee on Finance and Public Administration 1990, p. 86). The review led to the Government endorsing, in 1988, the first whole-of-government evaluation strategy.

Under the evaluation strategy, agencies were responsible for determining evaluation priorities and doing evaluations, but there were also formal requirements:

- all programs were to be evaluated every three to five years
- NPPs had to include evaluation plans
- forward evaluation schedules had to be submitted to the DoF
- evaluations had to be published (unless there were policy sensitivities, national security or commercial-in-confidence considerations).

The DoF gained an oversight role, providing advice, support and training in evaluation. Evaluation units were set up in most departments.

However, in 1991, a report by the Auditor-General found shortcomings in the quality and lodgement of portfolio evaluation plans, and suggested that the DoF was not doing enough to help departments develop their commitment to, and their capacity to undertake, evaluation. Other challenges at that time included a lack of skilled evaluators, difficulties obtaining and protecting evaluation budgets, lack of data, and difficulties in maintaining independence of evaluation units (Ryan 2003).

By the mid-1990s there were signs of improvement. The Auditor-General reported departments developing satisfactory evaluation plans, and there were mechanisms in place to improve the quality of evaluations (Auditor-General 1997, p. xiv). About three-quarters of all evaluations were made public and, according to one historian:

... there [was] clear evidence that evaluations were used intensively in the budget process: they provided a substantial contribution to the development of policy options and their consideration by the Cabinet'. (Mackay 2011b, p. 11)

Between the 1991-92 and 1994-95 Budgets, the rate of NPPs influenced by evaluation increased from 23 to 77 per cent (Department of Finance 1994b). Executive-level involvement⁸ in evaluation was commonplace and policies were in place for the endorsement and monitoring of recommendations.

However, issues persisted with departments obtaining staff with specific skills and appropriate training. According to the Australian National Audit Office (ANAO), in-depth investigations and departmental surveys revealed that ongoing performance measurement and reporting systems were influencing the conduct and planning of evaluations, but evaluation was not influencing the development or refinement of performance measures (Auditor-General 1997). While evaluation had advanced under the strategy, overall performance measurement had not. Views about the extent to which the strategy was successful in promoting good quality and useful evaluation across the APS remain mixed (appendix C).

From 1996, there was a shift away from a centralised approach. The DoF was no longer involved in monitoring evaluation, and evaluation plans for NPPs ceased. The formal evaluation strategy was discontinued in 1997, replaced by a less prescriptive, principles-based approach. Department heads were given greater discretion to determine what to evaluate, when and how.

In 1999, an Outcomes and Outputs Framework was introduced. The Framework was designed to place more emphasis on impact and the combined effects of activities (Podger 2018, p. 109). The Framework viewed performance through three questions — what does the Government want to achieve (outcomes), how does it reach those achievements (outputs), and how does it know it is succeeding (indicators)? (McDonald 2009, p. 4)

Further changes were made to the Government's performance framework in 2007. The Outcomes and Outputs Framework was replaced with a new Outcomes and Programs

⁸ Executive may include the Secretary, Deputy Secretary and/or First Assistance Secretary or equivalent.

Reporting Framework in 2009, a move that brought back program outcomes to performance reporting, but ultimately yielded no change in the use of evaluation in budget processes or publication of evaluations (Gray and Bray 2019, p. 12).

In 2013, public sector governance reforms led to the *Public Governance, Performance and Accountability Act 2013* (Cwlth) (PGPA Act). The Act updated performance reporting for Commonwealth entities, rolling requirements under the previous *Financial Management and Accountability Act 1997* (Cwlth) and the *Commonwealth Authorities and Companies Act 1997* (Cwlth) under the one umbrella. The PGPA Act provides a set of principles by which an agency can demonstrate the performance of its operations, while retaining flexibility in how it delivers that information. Evaluation was cited as one of the motivations for the creation of the Act:

The [PGPA] Bill strengthens and simplifies accountability across Commonwealth entities. It places a greater emphasis on performance monitoring, evaluation and reporting. It also provides greater clarification in relation to annual reporting. (Wong 2013, p. 15)

2.2 The current and changing landscape of evaluation

Responsibility for evaluation remains largely devolved ...

The Australian Government's performance reporting framework — the Enhanced Commonwealth Performance Framework — is the main instrument for measuring and reporting the performance of government activities (box 2.2). The framework underpins the PGPA Act (and its associated Rule).

There are no formal requirements for evaluation under the PGPA Act, however, the explanatory memorandum states that the Act seeks 'to link the key elements of resource management so that there is a clear cycle of planning, measuring, evaluating and reporting of results to the Parliament, Ministers and the public' (Wong 2013, p. 7). Australian Government agencies are required to report, through Annual Performance Statements, 'on the results actually achieved against the measures set out in their corporate plan and Portfolio Budget Statements, and on the basis of any reviews and evaluations they have committed to undertaking to demonstrate their performance in achieving their purposes' (Department of Finance 2017, p. 8).

Box 2.2 The Enhanced Commonwealth Performance Framework

The Enhanced Commonwealth Performance Framework came into force under the *Public Governance, Performance and Accountability Act 2013* (Cwlth) in July 2015. The framework represented a move to shift performance measures beyond the presentation of inputs and outputs to address directly whether Commonwealth entities were achieving their stated objectives (Morton and Cook 2015).

Under the framework, the overall performance of a Commonwealth entity or company is articulated across three documents.

1. **Portfolio Budget Statements** (PBS) that tie the allocation of resources to identified outcomes. PBS documents are updated throughout the year.
2. A **Corporate Plan** that sets out the objectives, capabilities and intended results (in line with the PBS) over a period of at least four years. In effect, the Corporate Plan is the primary document for articulating a Commonwealth entity or company's planned non-financial outcomes and how it intends to measure these outcomes.
3. An **Annual Performance Statement** that sets out the outcomes from commitments made in the Corporate Plan and PBS. Annual Performance Statements are included in the Annual Report of each Commonwealth entity or company, and (assuming evaluation is being used to support the agency's performance story) provides a platform to communicate the results of evaluations undertaken that year.

The flexibility written into the PGPA Act means evaluation planning and conduct is currently devolved to Australian Government agencies so they can decide what to evaluate, when, how, and for what purpose. As the Department of Social Services said, under the PGPA Act:

Agencies are responsible for planning and managing evaluations of their programs, but may use discretion to determine whether it is best to conduct an evaluation or other assessment of performance. (DSS 2018, p. 5)

There is also no centralised or whole-of-government evaluation strategy, policy or unit to oversee or guide the evaluation efforts of Australian Government agencies. As such, the extent to which evaluations are undertaken, and the support provided for evaluation, varies across agencies (section 2.2).

... but recent reviews have recommended some changes

Two recent reviews have looked at the state of evaluation in the APS.

The first — the *Independent Review into the Operation of the PGPA Act and Rule* undertaken in 2018 — commented on the decline in evaluation of Australian Government services and programs since the 1990s, and recommended that:

The Secretaries Board should take initiatives to improve the quality of performance reporting, including through more effective and informed use of evaluation, focusing on strategies to

improve the way Commonwealth entities measure the impact of government programs. (Alexander and Thodey 2018, p. 3)

The Independent Review of the APS, which was released at the end of 2019, also noted the long-term decline of in-house research and evaluation culture and capabilities within the APS and argued that evaluation capability and practices needed to be rebuilt.

The APS must provide successive governments with rigorous and innovative advice as the policy-making environment becomes increasingly complex and challenging. This will require a much stronger focus on research and evaluation in order to identify emerging issues and evaluate what works and why. (DPMC 2019e, p. 220)

Research for the review described the ‘approach to evaluation as ‘piecemeal’ (in both scope and quality) across the APS’. It also said that this approach ‘diminishes accountability and is a significant barrier to evidence-based policy-making’ (DPMC 2019e, p. 221).

The Review, with the aim of embedding a culture of evaluation and learning, recommended a central evaluation function be established (box 2.3). This function would:

- provide guidance and support for agencies on best-practice evaluation approaches
- drive a service-wide approach to evaluation and to uphold minimum standards of evaluation (with the main responsibility for evaluation continuing to reside with individual agencies)
- develop (for the Government’s consideration) a new strategic approach to evaluate past, present and proposed policies and programs (noting that the ad hoc nature of current processes means that a large number of programs continue year after year with no requirement to be evaluated)
- provide advice on how best to embed mandatory requirements for evaluation in Cabinet processes and budget rules
- oversee the establishment of an evaluation profession.⁹

⁹ Under the professions model articulated by the Independent Review of the APS (recommendation 20), the design of each profession is to be tailored to the needs of the discipline and informed by future service-side needs.

Box 2.3 **Independent Review of the APS recommendation on embedding a culture of evaluation and learning**

Recommendation 26

Embed a culture of evaluation and learning from experience to underpin evidence-based policy and delivery.

- Department of Finance (Finance) develop, for Secretaries Board agreement, an Australian Public Service (APS) wide approach to build evaluation capability and ensure systematic evaluation of programs and policies.
- Finance establish a central enabling evaluation function to support APS evaluation practices and expertise.
- Agencies to establish in-house evaluation functions and annual plans, and publish evaluations, unless exempt by the Cabinet.
- Government amend Cabinet and budget requirements to establish a systematic approach for formal evaluations.

Implementation guidance

- APS wide approach to:
 - plan evaluations of present and proposed programs and policies (including spending, revenue and regulation), and
 - provide guidance and support for agencies in best practice, building capability and the effective use of external experts.
- Finance Minister and Secretaries of the Department of the Prime Minister and Cabinet (DPMC) and Finance to settle work program for deep dive evaluations on cross-cutting topics — such as distributional or regional impacts of policies or programs against desired outcomes.
- Finance and DPMC to ensure agencies meet new Cabinet and budget process requirements.
- Draw on and support existing evaluation capabilities and expertise in agencies as much as possible. Finance may support major evaluations.
- Establish an evaluation profession.

Source: DPMC (2019e, p. 223).

The review acknowledged that one of the challenges to reversing the decline in evaluation across the APS was the risk of failures being exposed (while also noting the counterpoint that success can be celebrated and trust in government can be enhanced by making available objective and rigorous assessments of performance). It also commented on excessive risk aversion within the APS, quoting research commissioned for the review which found that:

... departments and agencies are often more concerned with reputational risk, seeking to pre-empt or divert criticism rather than learning from experience and feedback. (DPMC 2019e, p. 225)

The Review's recommendations around research in the APS more generally refer to developing 'a culture of innovation and experimentation to underpin evidence-based policy and delivery' (DPMC 2019e, p. 226).

Both reports (but the Independent Review of the APS in particular) pointed to the need for some form of cross-agency coordination in areas such as planning and prioritising evaluation, as well as developing evaluation capabilities across the APS.

In its response to the Independent Review of the APS, the Australian Government said it would:

- establish a small team in the DoF (supported by the Secretaries Board) to help build evaluation expertise and practices, drawing on existing expertise
- task the DoF with developing guidance to ensure systematic evaluation takes place in line with existing performance frameworks. The Department of Finance will work with the Department of the Prime Minister and Cabinet to embed evaluation planning in NPPs and Regulation Impact Statements
- establish an evaluation profession.

The Australian Government also noted that:

- short of requiring Cabinet approval to exempt evaluations from publication, the Government supported *appropriate* publication of completed evaluations
- no change to the budget and Cabinet process was required as current arrangements (robust advice on implementation challenges and risk mitigation strategies) should include consideration of evaluation (DPMC 2019d, p. 22).

The Government is in the early stages of implementing the parts of the Independent Review of the APS's recommendations it has agreed to so there is still some uncertainty about how these changes will operate in practice. However, these changes will establish some central, whole-of-government functions to guide and support evaluation activity across the APS. The implementation of the reforms, however, has been put on hold as the APS responds to the COVID-19 pandemic.

Other jurisdictions also adopt a centralised approach

A number of Australian states and territories, and other countries also currently adopt a centralised approach to evaluation.

The Governments of New South Wales, Queensland, Western Australia and the Australian Capital Territory have central agency oversight of evaluation and evaluation strategies (table 2.1). The Northern Territory Government is currently developing an evaluation strategy (NT Government 2019, p. 11), and while the Victorian Government does not have an evaluation strategy, it maintains some centralised guidance mechanisms, including mandatory evaluation of some policies and programs.

Table 2.1 Central evaluation frameworks in the states and territories

State/territory	Responsible agency	Strategy details
New South Wales	Department of Premier and Cabinet	The New South Wales Government's <i>Program Evaluation Guidelines</i> set out good practice principles for program evaluation and an Evaluation Toolkit looks at how they can be applied in practice. Government agency service-delivery clusters are required to report on completed, ongoing and planned evaluation activity to Cabinet's Expenditure Review Committee each year. Each cluster is required to individually prepare an annual schedule which represents a prioritisation of evaluation activity across the cluster.
Western Australia	Department of Treasury, Program Evaluation Unit	The Western Australian Government's <i>Evaluation Guide</i> outlines evaluation principles, evaluation methodologies (with examples), and a five-step evaluation process that evaluators should follow. It also provides templates and advice on planning an evaluation. The <i>Evaluation Guide</i> is overseen by Western Australia's Treasury Program Evaluation Unit. The unit has the dual roles of providing evaluation services as well as encouraging the establishment and bedding down of evaluation as a core business function across the public sector. Western Australia also operates a Sunset Clause that applies to any program over \$5 million for any one year. Programs subject to the Sunset Clause have a set period of funding and an evaluation plan is required in the initial funding submission. Continuation of funding after the cessation date is subject to consideration by Government, which is to be informed by evaluation of the program.
Queensland	Department of Premier and Cabinet, Performance Unit	The Queensland Government's <i>Evaluation Guidelines</i> offer advice on types of evaluation, principles and priorities, and outlines a seven-step evaluation process for agencies to follow. The Guidelines also offer advice on planning and commissioning evaluations as well as building a program logic and advice on data collection techniques.
Australian Capital Territory	Office of the Chief Minister	The ACT government's <i>Evaluation and Policy Guidelines</i> is the oldest state or territory plan in operation. The Strategy is a high-level plan that includes evaluation in the policy and budget cycles and scaling methodology in a 10-step evaluation process. The policy pays particular attention to the linking of evaluations with performance reporting, fostering a culture of evaluation, building capability and communicating and applying the lessons learned from the evaluation process.
Victoria	Department of Treasury and Finance	The Victorian Government does not have specific evaluation guidelines, but all lapsing programs seeking further funding are subject to some sort of mandated review – as articulated in Victoria's <i>Resource Management Framework</i>. The extent of required review can vary depending on the nature of the program to be extended. Programs with funding of more than \$20 million require an evaluation with endorsement of scope and terms of reference by the Department of Treasury and Finance, with input from Premier and Cabinet. Programs with funding of less than \$20 million require either an evaluation, or 'appropriate evidence' of the program's efficiency and effectiveness. The framework sets out some further requirements such as independence, timing of report submissions and questions to be addressed.

Sources: ACT Chief Minister's Department (2010); Department of Treasury (WA) (2015); Department of Treasury and Finance (VIC) (2019); DPC (NSW) (2016b, 2016a); Treasury and Trade (QLD) (2014).

Canada and the United States have also adopted a centralised evaluation model, while in the United Kingdom, central support is available to help agencies undertake evaluation.

- In Canada, the Treasury Board of Canada Secretariat has responsibility for: leadership for performance measurement and evaluation functions; performing and overseeing resource alignment reviews and centrally led evaluations of departmental spending; making publicly available reports and summaries of centrally led evaluations in a timely manner; and working with departments to ensure that data collection and reporting processes are efficient.
- In the United States, the *Foundations for Evidence-Based Policymaking Act of 2018* (assented to by the President in January 2019), requires agency heads to appoint Evaluation Officers and to submit annual evaluation plans to the Office of Management and Budget, and Congress.
- The United Kingdom has a series of guides produced by HM Treasury to assist agencies to develop and undertake evaluations. These include the Magenta Book and associated supplements focusing chiefly on evaluation, as well as the Green and Aqua books that provide further guidance on evaluation, appraisal and analysis. There is also a ‘What Works Network’ of nine expert centres. The network aims to improve the way government and other public sector organisations create, share and use evidence in decision making (appendix D).

2.3 How do Australian Government agencies currently approach evaluation?

As discussed above, the devolved nature of current arrangements means that Australian Government agencies have discretion about how they approach evaluation. A consequence of this is that there are wide variations in how agencies plan for, resource, undertake and use evaluation. The Australian Evaluation Society commented that its members reported:

... contrasting responses in terms of resourcing, effort and commitment from entities. At one end, there are indications of agencies that have reduced their effort and investment in evaluation and performance reporting. At the other, there are cases of increased development in information technology and reporting architecture, increased resourcing to the evaluation function, and a clearer understanding of the role and linkages from evaluation practice through performance and information management, to achieving accountability via being able to tell a performance story. (sub. 49, p. 21)

Strategies and forward planning

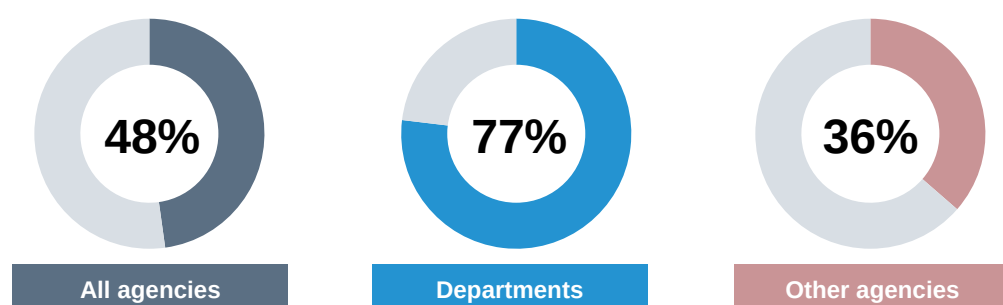
Evaluation strategies formally outline how agencies intend to approach evaluation. If well designed and followed, evaluation strategies can help agencies undertake evaluations in a consistent manner, and facilitate the strategic allocation of evaluation effort and resources to areas of greatest importance.

A number of Australian Government agencies have evaluation strategies or policies to guide how evaluation is planned for and undertaken within their agency. While less than half of all agencies responding to the Commission’s information request said they had an evaluation policy or strategy in place, most departments (77 per cent) reported having an active evaluation policy or strategy (figure 2.1). (Chapter 3 and appendix B contain more detail on the information request.)

Some strategies are publicly available while others are not. For example:

- the National Indigenous Australians Agency (NIAA), the former Department of the Environment, the former Department of Industry, Innovation and Science (DIIS), and the former Department of Infrastructure and Regional Development have evaluation strategies that are available on their websites
- the former Department of Agriculture and Water Resources, the Department of Health and the Department of Social Services have evaluation strategies but they are not available on their websites.¹⁰

Figure 2.1 Agencies with formal policies, frameworks, strategies, plans or guides for evaluation activity^a



^a Based on responses from 46 agencies: 13 departments and 33 other agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 1).

Most of the agencies’ evaluation strategies:

- outline a set of evaluation principles to follow
- outline a process for deciding on evaluation priorities within the agency (the processes used by agencies to set evaluation priorities are discussed further in chapter 6)
- require evaluation be considered when new programs are proposed
- provide guidance on how to develop and prepare a program logic.

¹⁰ Strategies and policies were provided before changes to Australian Government departments that came into effect in February 2020. Analysis of strategies and policies reflects the situation as at December 2019, prior to the changes taking effect.

Consistent evaluation principles across strategies

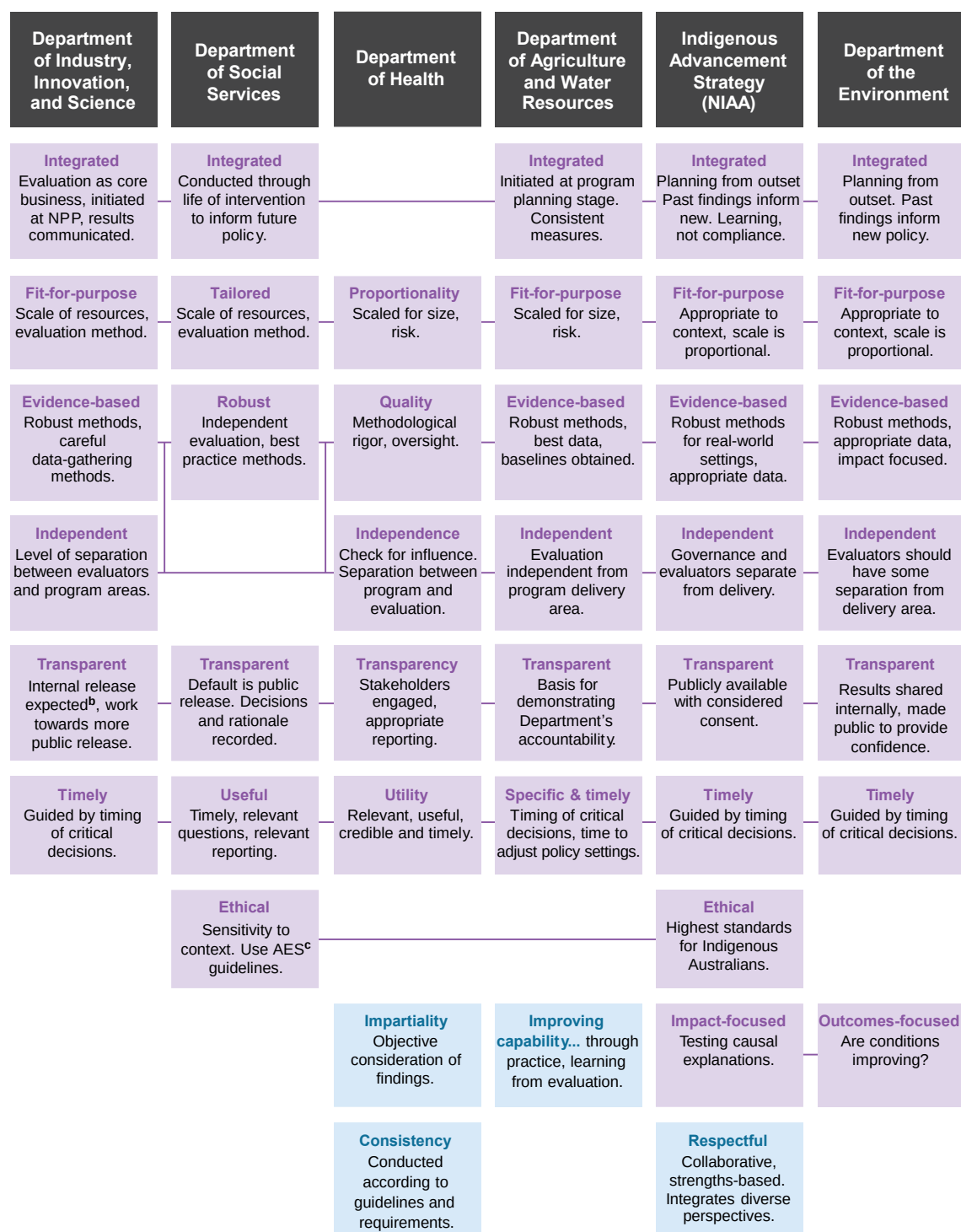
Evaluation strategies have similar principles to guide the conduct of evaluations (and reflect best practice principles for evaluation). Principles common to evaluation strategies (indicated in purple, figure 2.2), include:

- **Integrated** — evaluation is considered from the outset of planning for a program. The program is informed by previous evaluation findings.
- **Fit for purpose** — evaluation does not follow a ‘one-size-fits-all’ model. Evaluation activity is scaled for the size and risk profile of the program and appropriate evaluation and data methods are employed for the job at hand.
- **Evidence-based** — robust, best practice methods are used with careful data-gathering practices. Where possible, baseline data are considered and the results make sense in real-world situations.
- **Independent** — there exists a level of separation between those delivering a program and those who evaluate it.
- **Transparent** — evaluations are released internally within a department and preferably publicly.
- **Timely and useful** — evaluations take into account the needs of evaluation users and the timing of policy decisions.

Some participants to this project argued that additional principles (to the standard principles), are required for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. For example, the Victorian Community Controlled Health Organisation said:

It is important that the Productivity Commission recognise that current principles of evaluation, common across standard evaluation practice, do not provide enough information to work with Aboriginal people in a culturally appropriate way. A creative and open approach to [Indigenous Evaluation Strategy] principles is required, and presents genuine opportunities for collaboration and new perspectives from Aboriginal people and Communities. (sub 44, p. 7)

Figure 2.2 Strategies feature similar evaluation principles^a



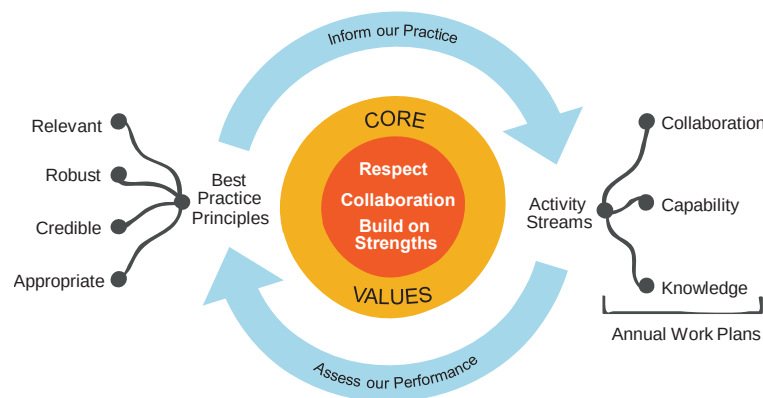
^a Departments appear as they did in December 2019, before Machinery of Government changes which came into effect in February 2020. ^b Where appropriate. ^c Australian Evaluation Society.

Sources: DAWR (2017); Department of the Environment (2015); DoH (2016); DIIS (2019a); DSS (2018); DPMC (2018b).

The NIAA's Indigenous Advancement Strategy (IAS) Evaluation Framework wraps its principles around a set of core values to recognise the strengths of Aboriginal and Torres Strait Islander people, communities and cultures. These values stipulate that all evaluations of IAS programs and activities need to test the extent to which they:

- build on strengths to make a positive contribution to the lives of current and future generations of Indigenous Australians
- are designed and delivered in collaboration with Indigenous Australians, ensuring diverse voices are heard and respected
- demonstrate cultural respect towards Indigenous Australians (figure 2.3).

Figure 2.3 IAS Evaluation Framework — core values and principles



Source: DPMC (2018b).

Guidance on ethical evaluation

Only some evaluation strategies (figure 2.2) include an ethical principle or provide guidance on how evaluation can be conducted ethically. That said, there are general frameworks for undertaking ethical research and specific ethical guidelines for conducting research with Aboriginal and Torres Strait Islander people and communities. For example:

- the National Health and Medical Research Council (NHMRC) has developed guidelines for researchers and stakeholders for ethical conduct in research with Aboriginal and Torres Strait Islander people and communities (NHMRC 2018a)
- the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) has *Guidelines for Ethical Research in Australian Indigenous Studies* (AIATSIS 2012) and is developing a new *Code of Ethics for Aboriginal and Torres Strait Islander Research* (AIATSIS 2019). The AIATSIS guidelines inform the NIAA's IAS Evaluation Framework (DPMC 2018b, p. 10).

These frameworks are informed by the recognition and respect for the rights of Indigenous peoples, as articulated in the United Nations Declaration on the Rights of Indigenous Peoples (box 2.4).

Box 2.4 Ethical conduct in research guidelines

AIATSIS Guidelines for Ethical Research in Australian Indigenous Studies

The Australian Institute of Aboriginal and Torres Strait Islander Studies guidelines emphasise the role of Indigenous people as full participants in research that involves their communities. The guidelines consist of 14 guiding principles grouped into six categories.

1. Rights, respect and recognition — including recognising diversity of Indigenous peoples, rights to self-determination and intangible heritage, respecting traditional knowledge, expressions, practices and innovations.
2. Negotiation, consultation, agreement and mutual understanding — including informed consent, ongoing consultation and negotiation achieving mutual understanding about proposed research, governed by a formal agreement.
3. Participation, collaboration and partnership — with full participation appropriate to people's skills and experiences in the project.
4. Benefits, outcomes and giving back — those involved in research should not be disadvantaged by it, with research results that respond to the needs and interests of Aboriginal and Torres Strait Islander people.
5. Managing research — plans for use, storage and guaranteed access to results.
6. Reporting and compliance — projects include mechanisms for reporting on compliance with these guidelines.

NHMRC Ethical conduct in research with Aboriginal and Torres Strait Islander Peoples and communities

NHMRC considers ethical research should improve the practice of all researchers working with Aboriginal and Torres Strait Islander people and communities, develop and strengthen the research capabilities of Indigenous people themselves and to enhance the rights of Aboriginal and Torres Strait Islander people as researchers, collaborators and participants. The guidelines are centred around six principles:

1. Spirit and integrity emphasises continuity between the present, past and future generations and the behaviours that hold Aboriginal and Torres Strait Islander values and cultures together. This first principle binds the next five.
2. Cultural continuity is maintaining the bonds between people and their environmental and spiritual domains, protecting spiritual and collective identities.
3. Equity is a commitment to showing fairness and justice that enables culture, history and status to be appreciated and respected.
4. Reciprocity, where ways of living and family relationships are kept strong through shared responsibility and obligation.
5. Respect is having regard for welfare, rights, knowledge, skills, beliefs, perceptions, customs and cultural heritage of the people involved in the research.
6. Responsibility includes caring for country, kinship bonds, caring for others and the maintenance of harmony and balance within and between the physical and spiritual realms.

Sources: AIATSIS (2012); NHMRC (2018a).

And specifically for ethics in evaluation:

- the Australian Evaluation Society (AES) makes it a requirement for membership that evaluators adhere to their code of ethics.¹¹ The code covers an evaluator's responsibility to the evaluation field, members of the public, the organisation and fellow members, and is set down in 18 principles
- BetterEvaluation, with funding from the NIAA, has produced a comprehensive set of principles on the ethical conduct of evaluation in Aboriginal and Torres Strait Islander settings¹². The principles are designed to act as a companion document to the AIATSIS Guidelines and the AES code of ethics.

Forward evaluation planning

Many agencies prepare (and a few agencies publish) forward evaluation plans. Good plans list the programs an agency is undertaking, the strategic priority of each program, the type(s) of evaluation that will be used and when each evaluation step will take place. Good examples of forward evaluation plans are provided on the websites of the Department of Foreign Affairs and Trade¹³ and the former DIIS¹⁴. The NIAA also has an evaluation work plan that is endorsed by an Indigenous Evaluation Committee (DPMC 2018b, p. 10).

Publishing a forward plan promotes transparency, scrutiny and trust of an agency's operations and performance. Developing an evaluation plan also provides an opportunity for agencies to consider how they will evaluate programs, and what data, capabilities and resources are required to undertake evaluations well.

Who undertakes evaluations?

About 30 per cent of Australian Government agencies and 60 per cent of departments that responded to the Commission's information request said they had an internal evaluation unit (figure 2.4).

Examples of agencies with an evaluation unit within their organisation structure include: the NIAA; the former DIIS; the Department of Foreign Affairs and Trade; the Department of Health; the Department of Social Services; the former Department of Infrastructure, Transport, Cities and Regional Development; and the former Department of Environment and Energy.

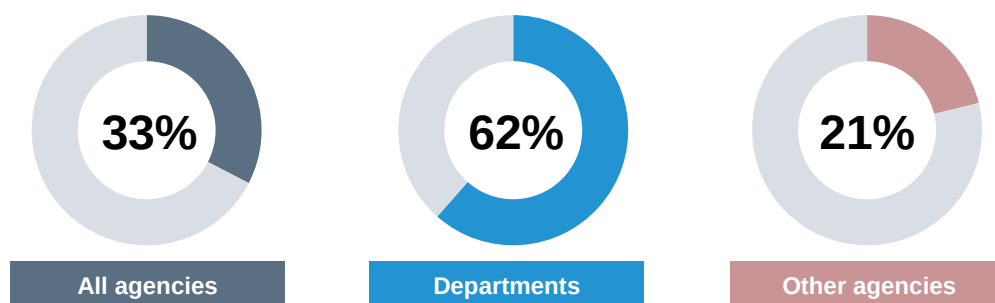
¹¹ https://www.aes.asn.au/images/stories/files/About/Documents%20-%20ongoing/code_of_ethics.pdf

¹² https://www.betterevaluation.org/sites/default/files/BetterEval_IndEval_Ethical_Protocol_v1_Aug2019.pdf

¹³ <https://www.dfat.gov.au/aid/how-we-measure-performance/ode/Pages/annual-aid-evaluation-plan.aspx>

¹⁴ <https://www.industry.gov.au/sites/default/files/2019-06/diis-evaluation-plan-2018-22.pdf>

Figure 2.4 **Agencies with a central evaluation unit^a**



^a Based on responses from 46 agencies: 13 departments and 33 other agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 4).

Most dedicated evaluation units:

- advise agency staff on the conduct of evaluations
- procure evaluation services from external providers
- review evaluations conducted within the agency
- collate evaluation evidence from across the agency.

Responses to the Commission's information request indicate that dedicated evaluation units have a range of roles. Just over half of evaluation units conduct evaluations (figure 2.5).

Resourcing for central evaluation units vary. For example, the evaluation unit within the former DIIS has 15 full-time equivalent employees while the central evaluation unit of another agency has a current annual budget of about \$190 000. It is common practice for Australian Government agencies to use external evaluators to evaluate policies and programs. Bray, Gray and t'Hart (2019, p. 17), in a background paper for the Independent Review of the APS, observed that 'a large amount of APS evaluation activity is undertaken by external evaluators'.

This accords with what the Commission was told in consultations for this project. It is also consistent with information provided in response to the Commission's information request — about 60 per cent of the evaluations provided to the Commission by the agencies that responded to the request were undertaken by an external evaluator.

Figure 2.5 Roles of central evaluation units^a



^a Based on responses from 15 agencies that reported having a central evaluation unit. Agencies were asked to select roles from a list. More than one response could be selected by each agency.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 4).

AusTender data between 2017 and 2019 also shows that on average, about 70 per cent (by dollar value) of tenders for evaluation services awarded by Australian Government departments each year were awarded to commercial consultants, and about 18 per cent were awarded to academic institutions. The remainder was made up by government agencies and non-government organisations.¹⁵

Reasons given by agencies for commissioning external evaluators (in consultations conducted by the Commission) included that:

- the agency did not have the capabilities needed to conduct the evaluation ‘in-house’ (which aligns with the findings of the Independent Review of the APS)
- using an external evaluator helps to provide a degree of independence between those evaluating the program and those responsible for its implementation.

Some agencies said they adopted a tiered approach, where they sought external evaluators to conduct evaluations of ‘flagship’ programs or programs considered to be of strategic importance, while evaluations of less important policies and programs were conducted by internal evaluation units.

¹⁵ Tenders in a given year for Australian Government departments were filtered for the word ‘evaluation’ in the description. Further manual filtering removed tenders that are obviously not program or policy evaluation, for example occupational performance evaluation. Lack of detail provided in the tender descriptions however meant results are indicative only.

The limitations of using external evaluators, however, were also pointed out by some agencies during consultations conducted by the Commission. Some of the comments were that:

- evaluation is a skill that needs to be practiced — individually and institutionally. If evaluations skills are not exercised, internal capability is diminished
- a lack of internal capacity in research, data and evaluation can hinder decision making within an agency
- evaluations — including when commissioned — need to be managed in a way to produce reports of publishable quality (and managing an external workforce comes with its own sets of challenges)
- the suitability of evaluators is not always visible in the tender process (there were reports of evaluations being delayed because agencies were required to up-skill consultants, including in cultural competency, during an evaluation)
- consultants have high charge-out rates, which restricts the time evaluators can spend on the ground
- consultant-produced evaluation reports can lack the consistency in methodology that an in-house evaluation unit can provide
- the nature of competitive tendering for evaluation contracts suggests the independence of an external evaluation is not necessarily assured.

The Independent Review of the APS also raised the issue of the use of contractors and consultancy services, noting that about a quarter of the submissions to the review commented on their use (DPMC 2019e, p. 185). There is a clear trade-off between leveraging the use of external contractors which can bring essential, sometimes niche expertise to a program area, and developing capacity within the APS.

Labour contractors and consultants are increasingly being used to perform work that has previously been core in-house capability, such as program management. Over the past five years, spending on contractors and consultants has significantly increased while spending on APS employee expenses has remained steady. (DPMC 2019e, p. 185)

Diminishing in-house practice and understanding of evaluation may affect the ability of agencies to procure and check the quality of externally-provided evaluation services and currently there are no mechanisms in place to gauge the level of capability in external procurement of services across the APS (DPMC 2019e, p. 264).

How are evaluations funded?

There is little formal, centrally-provided guidance on how evaluations should be funded in the APS. Consultations undertaken as part of this project pointed to the onus of securing resources to undertake evaluations laying primarily with agencies. Generally speaking, agencies said that funding for evaluating policies and programs is usually drawn from two sources (sometimes in combination):

-
- from the budget of the policy or program — for new policies and programs, this sometimes (but not always) includes funding allocated for evaluation through the NPP process
 - from a pool of funds ‘set aside’ by the agency for evaluating their policies and programs. This pool of funding is often sourced from the agency’s budget appropriation.

The Australian Evaluation Society said that, in their experience, funding evaluation through program and policy allocations was more common:

Feedback from AES members suggests that evaluation is more commonly funded from program budgets and new policy proposals, rather than a centralised budget. (sub. 49, p 22)

Both approaches have advantages:

- funding evaluation through a policy or program’s budget helps to ensure evaluation is considered at the planning and design stage
- the ‘pooled’ approach provides some flexibility. Because funding for evaluation is not tied to particular policies or programs, agencies can decide how to spend the budget based on where they consider there will be the greatest benefit. Pooled approaches are also more conducive to cross-cutting or strategic evaluations (that is, evaluations that extend beyond individual policies and programs).

However some agencies said that securing adequate funding for evaluations was not always easy. The Commission heard about funding for proposed evaluations being removed or reduced from NPPs before they were accepted. Some agencies indicated that while they aim for between five to ten per cent of each program budget for evaluation, the reality is that the proportion of program funding made available for evaluation is often significantly lower.

How do agencies share and use evaluation findings?

There is currently no requirement for Australian Government agencies to publish evaluations they undertake, nor are there any whole-of-government repositories or registers on which agencies can publish their evaluations.

For some agencies, the default practice is to make evaluation reports publicly available. The NIAA said that under the IAS:

... all evaluation reports or summaries are to be made publicly available. In cases where ethical confidentiality concerns or commercial in confidence requirements trigger a restricted release, summaries of the findings will be published in lieu of a full report. (sub. 86, p. 7)

The Department of Foreign Affairs and Trade has a default requirement for evaluation reports to be published, and maintains repositories of completed evaluation reports on its website (DFAT 2017a). The DIIS’s Evaluation Strategy also states that evaluation findings will be published where appropriate (DIIS 2019a, p. 39).

Most agencies said decisions about publishing evaluation reports were made on a case-by-case basis. And responses to the Commission information request indicates that a sizable proportion of completed evaluations are not being published. Of the 307 evaluations identified, only 43 per cent had a publically available final report. Publication rates vary between agencies with ten per cent of agencies reporting they ‘always’ published evaluation reports and about a third ‘very often’ while others publish relatively few (figure 2.6).

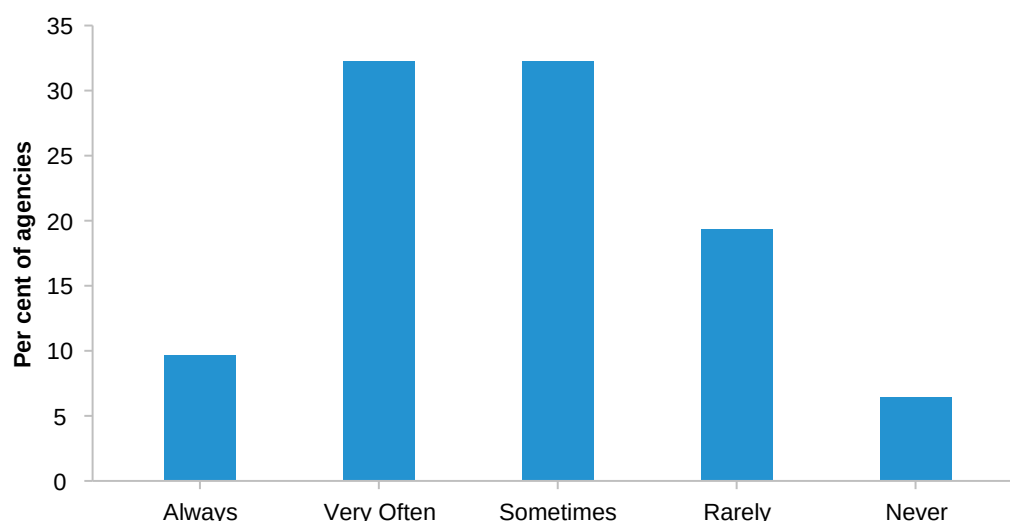
The Commission heard about a reluctance to publish evaluations when the findings were not positive and some suggested that requiring evaluations to be published could lead to pressure to remove or moderate findings. One participant also commented that:

Australian Government Departments ... can impose excessive controls over the release of program evaluation findings including restricting any information release whatsoever about the evaluation under way. The lack of guidance on these matters means that departmental staff interpret these restrictions in an ad hoc way and can be extremely restrictive and cautious depending on personalities. There have been situations where departmental staff have inserted clauses into contracts to restrict evaluators and peak Aboriginal representative bodies from sharing evaluation-related information. (Sophia Couzos, sub. 92, p. 4).

The publication of evaluations is discussed further in chapter 7.

Figure 2.6 Evaluation reports are not routinely published by default

How often are evaluation reports made available on agencies’ external websites?



^a Based on responses from 31 agencies. Sample excludes agencies that answered ‘not applicable’ to this question, which are typically agencies that report doing no recent evaluations.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 12).

On the use of evaluation, while there is strong agreement across Australian Government agencies that policy should be informed by evidence about what works, the Commission heard that in practice this is often not the case. One government agency told the Commission that a lot of program evaluations provide headline metrics, but not relevant findings and that ‘evaluation is often seen as a cost, not a benefit’. Similarly, others said that evaluations are often treated as compliance exercises and the focus is often on measuring outputs rather than impacts, outcomes or effectiveness. Some of the other themes were that:

- evaluation findings are often not timely enough to provide answers to policy questions
- evaluations can be of little value because they have not asked the right evaluation questions (chapters 4 and 5)
- funding and timeframes for evaluations can be inadequate and this can compromise the quality and usefulness of evaluation findings.

That said, some agencies pointed to evaluation findings that had influenced internal decision making and resulted in changes to policy. One example is the School Enrolment and Attendance Measure, which provided social work support to ensure children enrolled in and regularly attended school (Goldstein and Hiscox 2018, pp. 1–2). This program was stopped after several evaluations, including a randomised controlled trial, found no evidence that it was achieving its main objective of increasing school attendance (DPMC 2018a, p. 90).

Some agencies also require (and publish) management responses to evaluation findings, with examples including the former Department of Infrastructure, Transport, Cities and Regional Development and the former DIIS. DIIS’s Tradex Scheme offers customs duty and GST exemptions for importers who subsequently export their goods and was evaluated under the former DIIS’s evaluation strategy. The strategy requires management comments in response to evaluation findings attached to the final evaluation report, an example of which is provided in table 2.2. Such responses provide a useful and transparent way to demonstrate how evaluation findings are influencing decision-making and linked into the policy development and implementation process.

Table 2.2 Management response example
DIIS's Tradex Monitoring Evaluation Report 2018

<i>Recommendation</i>	<i>Response</i>	<i>Detail</i>
1. The benefits conferred to firms using the Tradex Scheme should continue to be available.	Trade and International supports this recommendation	
2. Investigate unification of duty concession mechanisms under one body.	Trade and International supports this recommendation.	This recommendation should be considered prior to agreeing any approach to the following recommendations. Although the report recommends an independent investigation, an effective outcome could be achieved by the Department and the Department of Home Affairs initiating a working group to consider administrative arrangements of Tradex and the Duty Drawback Scheme.
3. Improve the level of awareness and understanding of Tradex amongst Customs brokers and freight forwarders, and Customs officials.	Trade and International supports this recommendation.	Consideration of the effort and resources to improve awareness should be in line with potential benefits. Ensuring all outward facing documents are current and easily accessible is an initial step towards increasing knowledge of the scheme. Trade and International welcomes the finding that there is a high level of awareness of Tradex by Customs officials.
4. Modernise the application process.	Trade and International supports this recommendation.	If this recommendation is accepted, appropriate timeframes and performance indicators should be developed early to set deliverables
5. Review the key performance indicators and objectives of the scheme.	Trade and International supports this recommendation.	If this recommendation is accepted, appropriate timeframes and performance indicators should be developed early to set deliverables.

Source: DIIS (2019b).

3 The state of evaluation for policies affecting Indigenous people

Key points

- Participants' concerns about the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people were in four main areas:
 - the lack of evaluation and the ad hoc way that evaluations are undertaken
 - the quality of evaluations, including that methodologies adopted for evaluations of policies and programs are often not rigorous, and evaluation is not integrated into the policy cycle
 - the lack of engagement with Aboriginal and Torres Strait Islander people in determining what is evaluated, and in the design, conduct and interpretation of evaluations
 - the usefulness of evaluations, including that evaluations are often not effective in bringing about change and they tend to focus on policy- or program-specific questions rather than cross-agency or cross-topic lessons.
- The Commission sent an information request to Australian Government agencies asking about their evaluation practices and to provide copies of (or links to) evaluations published since 2016-17 and details on any unpublished evaluations. Some insights from evaluations provided to the Commission were that:
 - health was the area of policy with the highest number of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, followed by school education, labour and employment, and community and environment services
 - one-third of mainstream evaluations mentioned or provided results showing outcomes for, or effects of, mainstream programs on Aboriginal and Torres Strait Islander people
 - while most of the evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people attempted to measure impact, many had limited data. Most evaluations used both quantitative and qualitative methods, but few included cost-benefit analysis
 - Aboriginal and Torres Strait Islander people had minimal input on the planning, conduct or reporting on evaluations undertaken by Australian Government agencies that mentioned or provided results for Aboriginal and Torres Strait Islander people.
- A review of the evaluation evidence base for policies and programs affecting Aboriginal and Torres Strait Islander people found that there are significant gaps in the evidence about the impact of policies and programs and how outcomes could be improved. That said, some themes from the evidence are that: policies and programs implemented with an evidence base behind them are more likely to be successful; planning for evaluation and engaging with Aboriginal and Torres Strait Islander people and communities are critical for quality evaluation; and policies and programs need to be flexible for local contexts.

This chapter looks at the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. The chapter seeks to answer the following questions:

- what did participants say about the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people (section 3.1)?
- what do we know about recent evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people undertaken by Australian Government agencies (section 3.2)?
- what insights do we have from evaluation findings about what works or does not work for policies and programs affecting Aboriginal and Torres Strait Islander people (section 3.3)?

3.1 What we heard about the current state of evaluation

Participants' comments on the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people were in four broad areas. They included concerns about — the extent and ad hoc nature of evaluation; the quality of evaluations; the lack of engagement with Aboriginal and Torres Strait Islander people; and the limited usefulness of evaluations.

Concerns about the extent and ad hoc nature of evaluation

Some participants said that very few policies and programs affecting Aboriginal and Torres Strait Islander people are evaluated (so the evidence about what works or does not work is scant). Cape York Institute, for example, said:

... there is a lack of evaluation, where evaluations are conducted they are often not very useful in terms of driving change, and evaluations of interventions that show clear and convincing outcomes in Indigenous affairs are few and far between. ... evidence about 'what works', including for whom, under what circumstances, at what cost and why, remains scant. ... Given that the total spend on Indigenous Affairs in Australia is nearing \$35 billion annually it seems perverse and a measure of the problems, that under the current 'system' monitoring and evaluation does not improve results. (sub. 69, pp. 3,12)

And Children's Ground said:

Eleven years of Closing the Gap data is evidence of failed approaches and a failed system. This has been perpetuated by the lack of evaluation of key Indigenous policies and programs which means we do not have a national evidence base (sub. 73, p. 3).

The lack of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people has also been pointed out by others.

- A 2016 report that mapped the number of Indigenous-specific programs (Hudson 2016), found that just 8 per cent¹⁶ of Indigenous-specific programs had been evaluated (either during or after implementation).
- The Commission's *National Indigenous Reform Agreement Performance Assessment 2013-14*, concluded that while there was extensive reporting on outcomes for Aboriginal and Torres Strait Islander people, formal rigorous evaluations of policies and programs (both mainstream and Indigenous-specific) were 'relatively scarce' (PC 2015).

A recent Queensland Productivity Commission report on *Service Delivery in Remote and Discrete Aboriginal and Torres Strait Islander Communities*, however, commented that:

... the problem is not simply one of 'not enough evaluation'. It appears that there is a significant amount of time, energy and money ... spent on monitoring and evaluation in Indigenous affairs. For example, the Australian Institute of Health and Welfare's Closing the Gap Clearinghouse Research and Evaluation Register reveals 1 249 evaluation studies. Our tentative view is that the effort spent on evaluation in the Indigenous space is comparable to the effort spent evaluating social services, generally. ... A big part of the problem is that current evaluation efforts focus on the wrong things. Most effort appears to be on either ensuring accountability for expenditures (compliance) or demonstrating value for money, rather than on finding ways to improve outcomes. (QPC 2017, p. 227)

Other participants expressed concern about which policies and programs were evaluated and the ad hoc way evaluation decisions seem to be made. The Australian Institute of Aboriginal and Torres Strait Islander Studies, for example, said:

Evaluation of policies and programs and local outcomes have been undertaken in an ad hoc way, constrained by the parameters of a particular program or activity, such that there is no overarching logic that provides a picture of the impact that government is having in Aboriginal and Torres Strait Islander peoples' lives. (sub. 72, p. 4)

As discussed in chapter 2, there is no coordination of evaluation at the whole-of-government level, including for setting evaluation priorities (although some agencies have their own criteria for what to evaluate).

¹⁶ The study identified 1082 Indigenous-specific programs — 49 Australian Government programs, 236 state or territory programs and 797 programs delivered by non-government organisation (funded in part or full by government). Just 88 of the 1082 had been evaluated.

Concerns about the quality of evaluations

Participants also raised concerns about the methodologies adopted for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, including that evaluations often lacked rigour:

Too often, evaluations of policies affecting First Nations peoples have been poorly undertaken. ... Too often evaluations have been designed in such a way that causal inference is impossible. (Tony Dreise et al., sub. 33, p. 8)

[Aboriginal and Torres Strait Islander] agencies are often told the evaluator will use an already much rolled out methodology which may or may not be tweaked for the different settings and so it is blunt. It can be a ‘tick and flick’, ‘tick a box’ or one that is cheap for the evaluator to roll out repeatedly with minimal reflective practice on its appropriateness to the agency or community setting. (Liz Curran, sub. 2, p. 2)

A number of participants argued that evaluation is often an afterthought, rather than being integrated into the policy cycle, and this affects the quality of evaluations (including because the data needed for evaluation are not collected). For example, the National Aboriginal Community Controlled Health Organisation, said:

... the evaluation of services affecting Aboriginal and Torres Strait Islander people is in urgent need of reform. There is little evidence that evaluation is being used in the Indigenous Affairs policy cycle. To the extent it occurs, it appears to be done on an ad hoc basis without any strategy, and much of it seems to be ignored. (sub. 95, p. 1)

And the Victorian Aboriginal Child Care Agency:

Planning for evaluation has generally not been done by government funding agencies during the design and development of programs. Common practice has been that programs commence implementation with planning for an evaluation following, and requirements from funding agencies for participation in evaluation after the program has begun and often after the evaluation has been designed and the external evaluator engaged. Recognising the importance of data collection and the critical role of practitioners in contributing to evaluation, if evaluation data collection is not planned at the start of a program’s inception, there can be resistance to data collection, contamination of data and other implementation challenges. (sub. 26, p. 6)

Concerns about the quality of evaluations for policies and programs affecting Aboriginal and Torres Strait Islander people, including because evaluation is not integrated into the policy cycle, are not new. At the Commission’s 2012 Roundtable on *Better Indigenous Policies: The Role of Evaluation*, Deborah Cobb-Clark said:

Unfortunately, our current evaluation system generally produces poor-quality evaluations that in the end do not tell us very much. (2013, p. 90)

And Matthew James, commenting on the importance of building evaluation into policy design, said:

If evaluation and monitoring are not built into the policy development process some types of evaluation will not be possible and important data may not be collected. (2013, p. 117)

Lack of engagement with Aboriginal and Torres Strait Islander people

The lack of engagement with Aboriginal and Torres Strait Islander people on evaluations and evaluation priorities was a key area of concern for participants. Liz Curran, for example, said:

... [Aboriginal and Torres Strait Islander] peoples (First Nations Peoples) often know what works and why or why not and what the solutions and proactive steps will be for the community, but they are often not asked, listened to or this gets lost in translation and implementation if they are heard at all. (sub. 2, p. 1)

AbSec – NSW Child, Family and Community Peak Aboriginal Corporation also said:

In our experience, government agencies focus on external evidence and impose external solutions on Aboriginal communities, without sufficient regard for their appropriateness and limited evaluation of their effectiveness for local Aboriginal communities. Where evaluations have occurred, they have focused on the priorities and perspectives of government agencies, rather than those of Aboriginal communities and service users. In short, the intended beneficiaries of Aboriginal policy and programs tend to be marginalised in the design, delivery and evaluation of those programs. (sub. 9, p. 4)

The Queensland Indigenous Family Violence Legal Service said that the current ‘top down’ approach meant findings were of little or no benefit to Aboriginal and Torres Strait Islander people and communities (sub. 25, p. 4). Similarly, the Victorian Aboriginal Child Care Agency said:

The lack of engagement of Aboriginal organisations and communities in determining priorities has several impacts, including evaluations not focusing on issues of enough importance to Aboriginal communities, and data collection burden on Aboriginal organisations. (sub. 26, p. 6)

The recent Independent review of the Australian Public Service made a similar observation:

This is a damning assessment of years of focus on gaps and problems, not on strengths and assets. Too often, this approach has seen the APS do things to, not with, Aboriginal and Torres Strait Islander communities, and substantially fail to improve social and economic outcomes. (DPMC 2019e, p. 130).

As discussed in chapter 1, the key to good policy outcomes is putting the people who are intended to benefit from the policy at the centre of policy development and evaluation.

Evaluation findings not influencing policies and programs

The other area of concern for participants was about how evaluation findings were used, with some suggesting that findings were often not used to achieve meaningful policy change. Some also raised concerns about the lack of transparency of evaluation findings while others

said that evaluation tended to focus on policy- or program-specific questions rather than cross-agency or cross-topic lessons. Empowered Communities, for example, said:

... there are a significant number of policy and program areas whose impact remains unknown, with little transparency. (sub. 41, p. 10)

And Social Ventures Australia said:

We recognise that there is both a lack of evidence underpinning the majority of government funded programs affecting Aboriginal and Torres Strait Islander peoples, and a lack of transparency over which programs are backed by evidence and/or evaluations. (sub. 83, p. 11)

3.2 What policies and programs have been evaluated?

An information request for some insights

To get a better understanding of the extent and coverage of evaluations of Australian Government policies and programs affecting Aboriginal and Torres Strait Islander people, the Commission sent an information request to 182 Australian Government agencies seeking information on agency's evaluation policies, strategies, practice and use. Agencies were also asked to provide copies of (or links to) their published evaluations from 2016-17 to 2019-20, as well as details on unpublished evaluations undertaken over the same time period. Information was sought on evaluations covering both mainstream and Indigenous-specific policies and programs.¹⁷ Further information on the information request and more detailed results can be found in appendix B.

Forty-six agencies responded to the information request, including 13 government departments,¹⁸ 14 medium to large mainstream agencies, two Indigenous-specific agencies and 17 small mainstream agencies. This was an overall response rate of 25 per cent. However, the response rate from departments (excluding the three parliamentary departments) was 76 per cent, while the response rate for other agencies was 20 per cent.

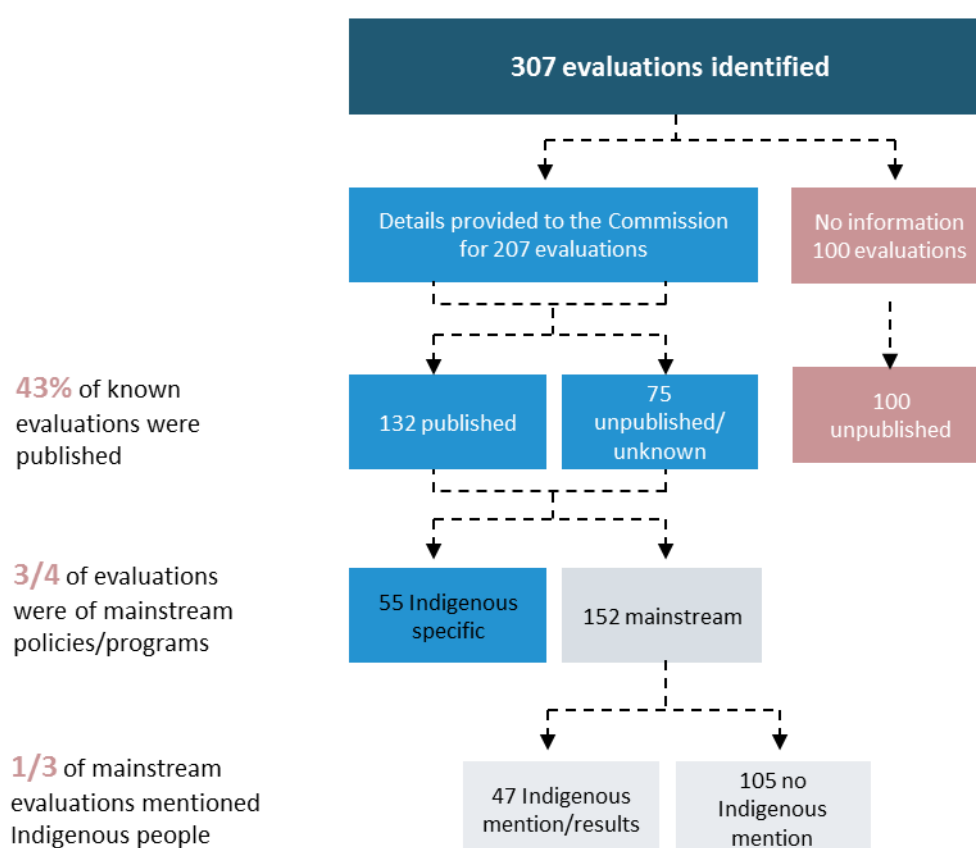
Agencies provided information on 207 evaluations conducted between 2016-17 and 2019-20. Two-thirds of these evaluations were published. Agencies also identified (but did not provide details on) a further 100 unpublished evaluations, giving a total of 307

¹⁷ Some caution needs to be exercised when drawing conclusions from agencies' responses to the information request as 1) not all agencies responded, 2) some agencies said they provided a subset of evaluations undertaken, 3) the assessment of evaluation conduct is based on information provided in the evaluation reports, and 4) there may be a response bias (those agencies with better evaluation practice, and/or larger and better resourced agencies may have been more likely to respond to the information request). More information on data limitations is included in appendix B.

¹⁸ The Department of Foreign Affairs and Trade is very active in evaluating its programs but is not included in the analysis in this chapter. The Department's evaluations focus on international development and assistance programs that do not affect outcomes for Aboriginal and Torres Strait Islander people. Chapter 2 includes some information on the Department's evaluation policies.

evaluations (figure 3.1). Overall, 43 per cent of Australian Government evaluations identified by the Commission through the information request were published. Publication rates varied between agencies. Ten per cent of agencies said they ‘always’ published evaluation reports and about a third ‘very often’, while others published relatively few (appendix B).

Figure 3.1 **Evaluations undertaken by Australian Government agencies between 2016-17 and 2019-20^a**



^a Counts represent information provided by the 46 agencies out of 182 agencies that responded to the Commission’s information request (appendix B has more information on response rates). Some agencies that did respond were not able to provide information on all evaluations done by the agency. The nature and size of policies and programs and evaluations varied greatly, therefore, counts only provide an indicative picture of overall evaluation practice. Agencies provided details for 207 evaluations. A further 100 unpublished evaluations were identified by agencies but no information on these was provided.

Data source: Productivity Commission Indigenous Evaluation Strategy information request.

Evaluation effort varies across agencies and topic areas

The responses to the information request show that evaluation effort varies across agencies — some agencies have undertaken a large numbers of evaluations in recent years while others have undertaken very few (appendix B). While it might be expected that smaller agencies do less evaluation than larger agencies, the responses to the information request shows that some (but not all) small agencies¹⁹ had been active evaluators in recent years, while some large agencies had undertaken very little evaluation.

Some agencies were active in evaluating a range of their mainstream policies and programs, however, only around one third of mainstream evaluations mentioned or provided results showing outcomes for, or effects of, mainstream programs on Aboriginal and Torres Strait Islander people.

The Commission looked specifically at evaluations provided by agencies that mentioned or provide results for Aboriginal and Torres Strait Islander people. Of these, there were 55 evaluations of Indigenous-specific policies or programs and 47 evaluations of mainstream policies or programs (figure 3.1).

The numbers of reported evaluations also varied across service areas.

- Most evaluations (one quarter) of Australian Government evaluations of policies and programs mentioning or providing results on Aboriginal and Torres Strait Islander people provided to the Commission were for health services, followed by school education, labour and employment, and community and environment services.
- Areas of significant government expenditure for policies and programs affecting Aboriginal and Torres Strait Islander people where there were few or no recent evaluations (based on those provided to the Commission) included social security payments and healthcare subsidies and support (including Medicare and the Pharmaceutical Benefits Scheme, box 3.1).²⁰

Notably, very few evaluations mentioning Aboriginal and Torres Strait Islander people examined issues cutting across agencies or service areas — only three Indigenous specific policy evaluations and one mainstream program evaluation covered multiple agencies or service areas.

Fewer than 10 per cent of Australian Government evaluation reports that mentioned or provided results for Aboriginal and Torres Strait Islander people included information on how the Government or the commissioning agency had used the evaluation findings or recommendations.

¹⁹ The low response rate of small agencies to the Commission's information request means that it is not possible to generalise about evaluation practice in small and medium agencies. It is also not possible to generalise about evaluation practice in small Indigenous specific agencies.

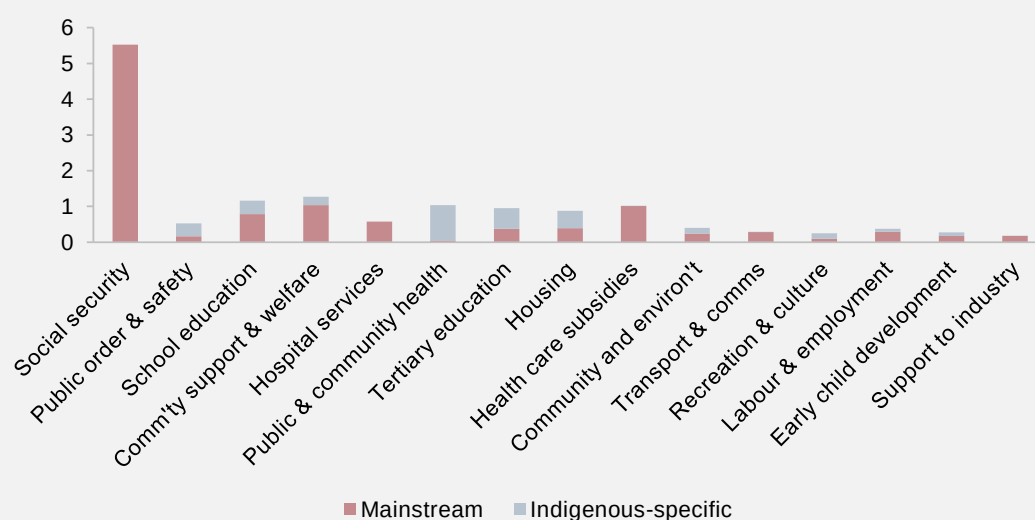
²⁰ There may have been evaluations in these areas (and others) that were not provided to the Commission, or evaluations were undertaken by entities other than Australian Government agencies. Policies and programs may also been evaluated in earlier years. For example, section 3.3 highlights a number of valuable lessons from evaluations in the area of housing.

Box 3.1 Comparing patterns of evaluation and expenditure

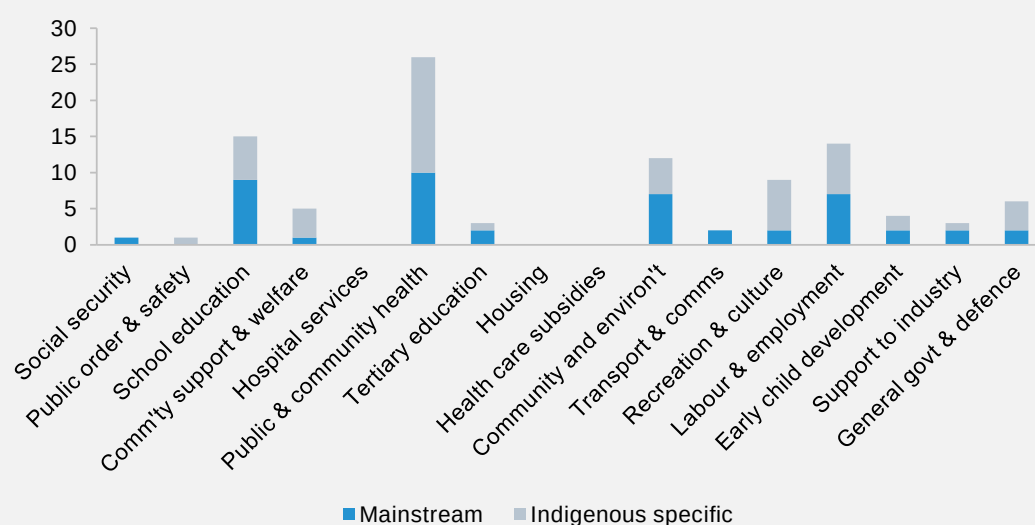
Data from the Indigenous Expenditure Report (SCRGSP 2017) show the distribution of Australian Government direct expenditure on Aboriginal and Torres Strait Islander people by service area. The areas attracting the most Australian Government direct Indigenous expenditure in 2015-16 (the most recent data available) were social security (income support payments and pensions) (\$5.5 billion), public and community health services (\$1 billion), health care subsidies (Medicare and the Pharmaceutical Benefits Scheme) (\$1 billion) and tertiary education (\$1 billion).

As mentioned in chapter 1, the majority (75.5 per cent) of Australian Government expenditure on Aboriginal and Torres Strait Islander people in 2015-16 was for mainstream services (services available to all Australians) (SCRGSP 2017). The remaining 24.5 per cent was for Indigenous-specific services (services designed specifically for Aboriginal and Torres Strait Islander people)

Australian Government total Indigenous expenditure (mainstream and Indigenous specific) by service area (\$ billion)



Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people, by service area (number)



Source: SCRGSP (2017); Productivity Commission Indigenous Evaluation Strategy information request.

Mixed methods dominate

About two thirds of the evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people used a mix of qualitative and quantitative methods. About a quarter were based primarily on qualitative methods, such as interviews and focus groups with service users and service providers (chapter 4 discusses qualitative and quantitative methods in more detail). About one in seven evaluations only used quantitative methods. The most common evaluation methods used were:

- literature reviews and analysis of documents generated as part of the policy or program
- interviews, consultation meetings and focus groups with service providers, representative bodies, service users and subject matter experts
- analysis of administrative data
- surveys of service users or providers and other relevant groups.

Quantitative analysis in Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people mainly involved analysing administrative data and surveys. Some evaluations had good data on outcomes, however, the majority of quantitative analysis was for activities, outputs, and user perceptions and feedback.

Most evaluations attempted to measure impact, but most had limited data and/or did not have a control group. Very few include cost-benefit analysis. Chapter 9 provides more information on the use of data in evaluations (and how data collection can be improved).

Limited engagement with Aboriginal and Torres Strait Islander people

The responses to the Commission's information request are consistent with participants' concerns about a lack of engagement with Aboriginal and Torres Strait Islander people.

- Only about one-fifth of evaluations of Indigenous-specific policies and programs involved engaging with Aboriginal and Torres Strait Islander people in evaluation decision making.
- Just one of the 47 mainstream evaluations involved Aboriginal and Torres Strait Islander people in planning and decision making.

It is important to note, however, that responses to the information request showed that the lack of involvement of key stakeholders was not unique to evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. For evaluations more generally, most or all decisions about evaluation design and governance were made by government agencies or evaluation consultants without input from those affected (or likely to be affected), by the policy or program.

Responses to the information request showed that:

- about half of Indigenous specific-evaluations engaged with Aboriginal and Torres Strait Islander people as part of the research process
- about one third of mainstream evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people engaged with Aboriginal and Torres Strait Islander people as part of their research
- engagement with Aboriginal and Torres Strait Islander people in evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people was mostly as participants in consultation interviews or group discussions or as survey respondents
- seven of 55 Indigenous-specific evaluation reports mentioned having Aboriginal and Torres Strait Islander people as part of the research or evaluation team.

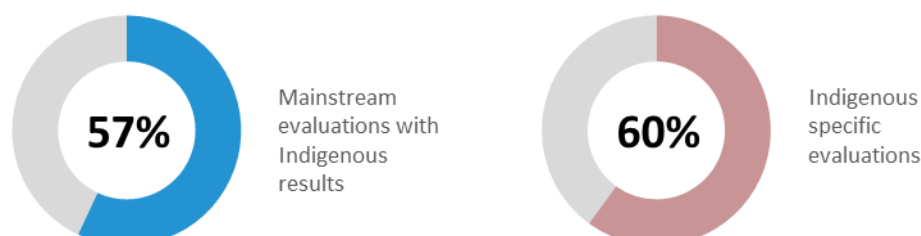
No mainstream evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people reported Aboriginal and Torres Strait Islander people being engaged in evaluation analysis or report authorship, or that the results were provided back to Aboriginal and Torres Strait Islander people. For Indigenous specific evaluations around one quarter engaged Aboriginal and Torres Strait Islander people in evaluation analysis or reporting and around one eighth provided results directly back to Aboriginal and Torres Strait Islander people.

The responses to the Commission's information request highlight the current inadequate level of engagement with Aboriginal and Torres Strait Islander people on evaluation with an obvious need for considerable improvement.

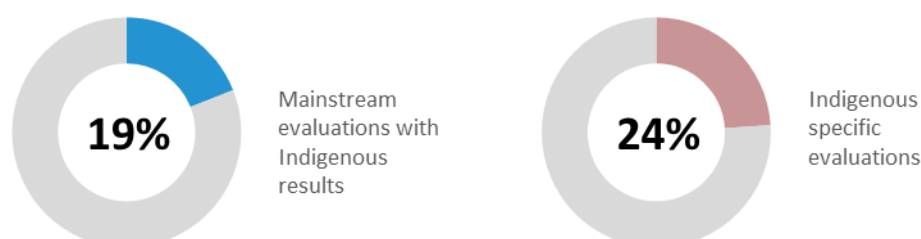
Figure 3.2 illustrates some characteristics of evaluations that mention or provide results for Aboriginal and Torres Strait Islander people.

Figure 3.2 Selected characteristics of evaluations that mentioned or included results for Aboriginal and Torres Strait Islander people^a

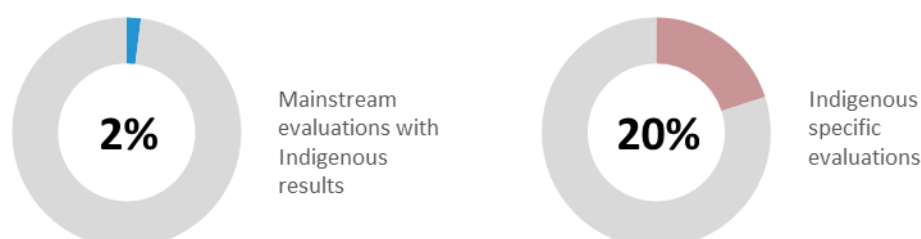
Around 6 out of 10 evaluations were done by external consultants



A minority of evaluations included formal ethics assessment



A minority of evaluations included Aboriginal and Torres Strait Islander people in decision-making



^a Counts represent information provided by the 46 agencies out of 182 agencies that responded to the Commission's information request. Some agencies that did respond were not able to provide information on all evaluations done by the agency. The nature and size of policies and programs and evaluations varied, therefore, counts only provide an indicative picture of overall evaluation practice.

Data source: Productivity Commission Indigenous Evaluation Strategy information request.

Australian Government evaluations are only part of the story

Evaluations by state and territory governments

Evaluations undertaken by state and territory governments of policies and programs affecting Aboriginal and Torres Strait Islander people also contribute to the evidence base on what works and what can improve outcomes for Aboriginal and Torres Strait Islander people. As

the Western Australian Government said, ‘State and Territory governments and local service providers are also key partners in policy implementation and evaluation’ (sub. 74, p. 1).

The New South Wales (NSW) Government has invested in recent years in its community focused plan for Aboriginal Affairs, OCHRE — which stands for opportunity, choice, healing, responsibility and empowerment. Through OCHRE, the NSW Government works with Aboriginal people and communities to support their social, cultural and economic aspirations. In 2018, the first stage of an independent evaluation of OCHRE was completed. Commenting on lessons learnt, Aboriginal Affairs NSW said:

... the practice and experience of the OCHRE evaluation in its first three years has led the NSW Coalition of Aboriginal Regional Alliances (NCARA) to call for a stronger embrace of interweaving evaluation, research, policy, programming and services delivery. ... weaving together Aboriginal ways of being, knowing and doing with Western knowledge threads ensures that policies, programs, and services are co-owned and co-produced by communities and that such threads possess both cultural integrity and public confidence. AANSW will continue to work with the NCARA to develop this methodology as the OCHRE evaluation proceeds through its remaining six years. Developed from practice, this approach has considerable promise in the evaluation of policies and programs affecting Aboriginal people. (sub. 70, p. 1)

The OCHRE evaluation involves extensive engagement with six communities/regions. One distinctive aspect of the OCHRE evaluation is that the decision about whether the evaluation report for each region/community is published is made by representatives from each region/community. Chapter 5 has more information on OCHRE and its evaluation.

Box 3.2 provides some examples of state and territory government evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

Box 3.2 **Some examples of State and Territory Government evaluations**

- *Practice First* — NSW Department of Family and Community Services — evaluation of a framework for practice introduced to improve the quality and effectiveness of child protection — a mixed methods mainstream evaluation with some results for Aboriginal and Torres Strait Islander people (Cortis et al. 2019).
- *48 Hour Follow Up* — NSW Department of Health — a mainstream mixed methods evaluation of outcomes for Aboriginal and Torres Strait Islander people in public hospitals (University of Newcastle 2016).
- *Banned Drinker Register in the Northern Territory* — NT Department of Health — a mainstream mixed method process evaluation including results for Aboriginal and Torres Strait Islander people (Smith and Adamson 2018).
- *Northern Territory Sobering Up Shelters* — NT Department of Health — a mixed method mainstream evaluation with results for Aboriginal and Torres Strait Islander people (PWC Indigenous Consulting 2018).
- *Cradle to Kinder* — Victorian Department of Health and Human Services — a mixed method mainstream evaluation including results for the Aboriginal component of the program (DHHS 2017).
- *Improving Cultural Responsiveness of Victorian Hospitals* — Victorian Department of Health and Human Services — an evaluation of the cultural responsiveness of Victorian mainstream public hospitals for Aboriginal people (DHHS 2016).
- *Aboriginal Health Case Management and Care* — Victorian Department of Health and Human Services — a qualitative Indigenous specific evaluation (Effective Change 2016).
- *Specialist Domestic and Family Violence Court Trial in Southport* — Queensland Department of Justice — a mainstream mixed methods evaluation with results for Aboriginal and Torres Strait Islander people (Bond et al. 2017).
- *Restorative Justice Project 12-Month Program Evaluation* — Queensland Department of Justice and Attorney-General — a mainstream mixed methods evaluation with results for Aboriginal and Torres Strait Islander people (Restorative Justice Evaluation Team 2018).
- *Blurred Borders Legal Communication Tools* — Legal Aid Western Australia — an Indigenous-specific qualitative process evaluation (Legal Aid WA 2019).
- *Early Childhood Schools and the Koori Preschool Program* — ACT Education Directorate — a mainstream mixed methods evaluation with results for Aboriginal and Torres Strait Islander people (Power et al. 2016).

Evaluations by Aboriginal and Torres Strait Islander organisations

Many Aboriginal and Torres Strait Islander communities and organisations value evidence to use in their work in delivering services and advocating for or informing decisions made by governments. The Central Land Council explained that:

For thousands of years Aboriginal people survived in the Australian landscape relying on their intricate knowledge of the land and its plants and animals. (CLC nd, p. 1)

That said, Aboriginal and Torres Strait Islander organisations face a number of constraints to undertaking formal evaluations. These include a high level of dependence on grants and service contracts with government agencies and the extent to which they can undertake evaluation depends on whether funding is provided for evaluation. However, consistent with the trend towards greater engagement with and conduct by Aboriginal and Torres Strait Islander people in evaluation, Aboriginal and Torres Strait Islander organisations are increasingly investing in building evaluation capability. The Victorian Aboriginal Child Care Agency, for example, said that in recent years it has:

... significantly increased its investment in evaluation and research with the development of a Research and Evidence Development Team. ... VACCA has recently been supported by federal and state government agencies to build internal systems and capability. This focus on building our capability supports self-determination, is more sustainable, ensures that evaluation priorities reflect those of Aboriginal communities and that approaches and methods are appropriate to the context. (sub 26, pp. 1, 5)

Indigenous Community Volunteers²¹ also reported that:

Over the past six years, ICV has invested in evaluation capacity building and the design of an inclusive, culturally safe and relevant evaluation approach to contribute to better outcomes for First Nations Peoples. (sub. 88, p. 2)

The Institute for Urban Indigenous Health (IUIH), an Indigenous community controlled health organisation in south-east Queensland, has commissioned several evaluations to inform and improve the work that it does (Nous Consulting 2019; SAHMRI 2016). The IUIH is leading the governance and commissioning of evaluation, engaging expert evaluators and researchers and drawing on a mix of qualitative and quantitative analysis.

Other sources of information on evaluation and related evidence

As noted in chapter 2, there is currently no central register of evaluations of Australian Government policies and programs in general or with a specific focus on the effects of policies and programs on Aboriginal and Torres Strait Islander people. However the Closing the Gap Clearinghouse — a Council of Australian Governments' initiative jointly funded by all Australian Governments — operated from 2009 to 2014.

The Clearinghouse operated as a website and included:

- 36 summary review papers that collated evidence from research and evaluation on what works to improve outcomes for Aboriginal and Torres Strait Islander people
- an assessed collection of evaluation and research providing information on what works to overcome Indigenous disadvantage
- a general collection of research and evaluation on topics related to the Closing the Gap building blocks but not assessed by the Clearinghouse

²¹ Indigenous Community Volunteers changed its name to Community First Development in early 2020.

-
- a research and evaluation register listing government research and evaluations relevant to the Closing the Gap building blocks.

The collections included Australian and international research and evaluation relating to both general populations as well as indigenous peoples (AIFS nd; AIHW 2017a).

The Clearinghouse web pages remained available after funding ceased in 2014. In 2017 the site was archived. The archived site is still available, although it is no longer possible to search the collection databases (AIHW 2017a). The 36 Clearinghouse papers reviewing the evidence across a range of topics are available on the Australian Institute of Health and Welfare (AIHW) and Australian Institute of Family Studies (AIFS) websites (AIFS nd; AIHW 2017a). Appendix C provides more information about the Clearinghouse.

Other information on Indigenous related evaluation

In 2010, the Office of Indigenous Policy Coordination in the Department of Families, Housing, Community Services and Indigenous Affairs published a *Directory of Commonwealth Government Evaluations, Audits and Reviews of Indigenous-specific and Relevant Mainstream Programs* (OIPC 2010). It listed 186 evaluations and performance audits both completed and planned across a range of Australian Government agencies. Twenty (11 per cent) of the evaluations and performance audits were for mainstream programs on Aboriginal and Torres Strait Islander people and 85 per cent of the evaluations and performance audits were for programs managed by three agencies — the Department of Families, Housing, Community Services and Indigenous Affairs, the Department of Health and Ageing and the Department of Education, Employment and Workplace Relations (the three agencies who managed the bulk of Indigenous specific programs at that time).

The successors to these agencies²² responding to the Commission's information request continue to contribute the majority of Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people in the period 2016-17 to 2019-20:

- 76 per cent of Indigenous-specific evaluations
- 60 per cent of mainstream evaluations
- 69 per cent of all evaluations.

²² In late 2019 and early 2020 when agencies were responding to the Commission's information request for this project, Department of Education and Training, Department of Employment, Skills, Small and Family Business, Department of Health, Department of Social Services and National Indigenous Australians Agency were the successors to the 2010 Department of Education, Employment and Workplace Relations, Department of Health and Ageing and Department of Families, Housing, Community Services and Indigenous Affairs. Further changes to some Australian Government agencies occurred in February 2020 but this was after returns to the information request had been completed by the relevant agencies.

A few agencies maintain centralised registers or lists of evaluation reports. For example:

- the Department of Foreign Affairs and Trade's Office of Development Effectiveness maintains an easy to use webpage providing links to a wide range of published evaluation reports
- links to published evaluation reports for the NIAA are available through published annual evaluation plans, which users can find after opening the evaluation plans
- the Department of Social Services maintains an internal list of evaluations for reference by its staff.

Other sources of research information and evidence include:

- the Australian Indigenous Health Infonet (<https://healthinfonet.ecu.edu.au/>)
- the Indigenous Justice Clearinghouse (<https://www.indigenousjustice.gov.au/>)
- the Lowitja Institute (<https://www.lowitja.org.au/>).

3.3 What does the evidence tell us about what works (or does not work)?

Evaluations are undertaken to gain insights into what works. So what lessons and themes have emerged from the evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people?

Much of what we have learned from evaluations relates to specific program and policy areas, although many findings have wider applicability. One of the most important lessons from evaluations — unsurprisingly — is that programs and policies implemented with an evidence base behind them are more likely to have successful outcomes (Bowes and Grace 2014). This does not imply that policy makers should shy away from experimenting in the absence of an evidence base, or that they should wait for evidence when an urgent response is required. Rather, it means that a stronger evidence base increases the chances of a successful policy outcome. And that all policy and program interventions should have sound logic behind them based on previous experience.

The need for better evaluation evidence in the Aboriginal and Torres Strait Islander policy context is highlighted by the persistence of wide gaps between Aboriginal and Torres Strait Islander people and non-Indigenous people across many indicators of wellbeing — life expectancy, health, educational attainment and employment — despite more than ten years of Closing the Gap initiatives (chapter 1).

Engaging Aboriginal and Torres Strait Islander people and communities is crucial

Evaluations of successful programs highlight the importance of engaging Aboriginal and Torres Strait Islander people and communities in the design, development and delivery of programs, including through providing training and resources. Programs also benefit from local knowledge (often facilitated by ‘bottom up’, as distinct from ‘top down’, approaches) (box 3.3).

The participation and engagement of Aboriginal and Torres Strait Islander communities is consistent with self-determination and helps ensure programs are consistent with Aboriginal and Torres Strait Islander community priorities and needs. It also ensures that Indigenous knowledges and cultural beliefs and practices are valued, while promoting trust and cultural capability. And it can assist in recognising existing strengths and building further capability in Aboriginal and Torres Strait Islander communities.

Policies and programs need to be flexible for local conditions

Another theme from evaluations is the need for flexible design and delivery within policies and programs to ensure local needs and contexts are considered, and to reflect differing local priorities. While there are common features that make programs successful, there is no ‘one size fits all’ policy — policies and programs need to be tailored for local needs, circumstances and community priorities. While scaling issues (that is, what works in an area of one size might not work in an area much larger or smaller) represent one element of this, needs and priorities can potentially differ significantly even between local areas that might otherwise seem quite similar in terms of culture, size and demographics.

Evaluation requires planning

Some of the key lessons from evaluations relate to how to successfully undertake evaluations.

- One lesson is that there are significant benefits from planning for evaluation in the formative stages of policy formation, including by making provision for funding for evaluation, and ensuring data is available for the evaluation.
- Another is that there are benefits from agencies having clear plans and strategies to routinely undertake research and evaluation to identify successful (and less successful) aspects of policies and programs and to build an evidence base (for accountability reasons, to ensure successful programs remain, and to have a basis for potentially strengthening, consolidating or expanding programs).

Box 3.3 Some learnings from evaluations have widespread application

Evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people have produced a number of lessons about 'what works' for successful policy and program development and implementation. These include:

- participation and engagement of Aboriginal and Torres Strait Islander communities in the design, development and delivery of policies and programs is needed to ensure they are consistent with Aboriginal and Torres Strait Islander community priorities. It is not enough to have policies and programs that succeed in meeting the Government objectives. The policies and programs are only worthwhile if they successfully meet objectives that are actually valued by the community
- there is a need for flexible design and delivery so local needs and contexts can be taken into account
- active local community participation and engagement in development and delivery (bottom-up approach) leads to better outcomes
- participation and engagement of Aboriginal and Torres Strait Islander people better ensures the valuing of Indigenous knowledges and cultural beliefs and practices
- a community-wide approach to policy setting and implementation is needed
- the need to build trust and engage in a culturally safe manner (including in ways that respond to the cultural, social and economic context of Aboriginal and Torres Strait Islander people). Trusting relationships and genuine, collaborative and sustainable partnerships are required
- objectives need to be clearly determined early to properly inform policy responses
- commitment is required from high level leadership of relevant organisations and agencies to address the determinants of Aboriginal and Torres Strait Islander wellbeing
- employing Aboriginal and Torres Strait Islander staff and ensuring they participate fully in program design, delivery and evaluation, including through provision of adequate training and resources where necessary, strengthens capability of Aboriginal and Torres Strait Islander staff
- there are benefits from developing committed, skilled staff (whether Indigenous or non-Indigenous) and providing diversity and cultural safety training
- where programs demonstrate success, providing adequate, sustainable resources to ensure long-term viability
- service delivery needs to be culturally safe (generally facilitated by cultural capability training and collaboration between Indigenous and non-Indigenous staff)
- adopting a strengths-based perspective builds and develops the existing strengths, skills and capacities of Aboriginal and Torres Strait Islander people
- services need to be coordinated, and the tendency of agencies and governments to operate in silos needs to be avoided.

Sources: Al-Yaman and Higgins (2011); Bowes and Grace (2014); Closing the Gap Clearinghouse (2013; 2013); Hunt (2016); Osborne, Baum and Brown (2013); QPC (2017); SCRGSP (2016); Ware (2013).

Lessons (and knowledge gaps) in individual service areas

While many valuable lessons from evaluations can be applied across various policy areas, a number of others relate specifically to policies and programs in specific service areas. Examples of these are discussed below for a selection of areas of public policy aimed at encouraging economic and social participation (education, health and housing) where, as discussed earlier in the chapter, there has been a higher level of evaluation activity than in other areas. In other areas of policy where there is less evaluation, the ‘knowledge gap’ is likely to be significantly larger.

Education

The evidence base for education policies and programs could be described as ‘mixed’. That said, there are still a number of themes coming through evaluations. They include:

- education practitioners and policy makers need to understand the importance of cultural factors in schooling and this needs to be reflected in the development of policies and programs
- strong leadership within schools is critical to delivering better results, including not being satisfied with being an ‘average’ school (there are benefits from setting high expectations for students, box 3.4)
- cultural competence and cultural safety needs to be addressed in program delivery
- respect and empathy for students produces better outcomes
- a positive Aboriginal and Torres Strait Islander identity should be promoted
- curriculum and pedagogy adopting culturally safe approaches, including Aboriginal and Torres Strait Islander languages, can promote increased engagement
- having a ‘whole of institution’ (for example, school or university) approach, involving senior management, to improving Aboriginal and Torres Strait Islander outcomes
- well-trained, high quality teachers are critical to success
- flexibility to adapt to the needs of the local community (especially where mainstream education programs are provided to Aboriginal and Torres Strait Islander people)
- Aboriginal and Torres Strait Islander communities should participate and be actively engaged in the planning and provision of education
- participation of parents in education is critical; building their skills also has benefits
- targeted skill development and mentoring programs are likely to benefit at risk students
- schools need to be flexible and focused on developing social capital
- effective use of data can assist in determining student abilities and the success of interventions by schools

- reading programs work best when they start early in a child’s education (AIHW and AIFS 2013; Al-Yaman and Higgins 2011; Bowes and Grace 2014; NCS 2012; Osborne, Baum and Brown 2013; PC 2016a).

Box 3.4 Success in remote schools: what works?

Remote and very remote schools face significant challenges getting outcomes for Aboriginal and Torres Strait Islander students that are commensurate with all students. The What Works team from National Curriculum Services undertook a study for the Australian Government in 2012 looking at 11 high performing remote schools.

While each of the 11 schools were located in distinctive contexts and had adopted strategies and actions in response to that context, there were some high-frequency practices that were across the 11 settings.

The study identified seven common ‘high-frequency themes’. These were:

1. leadership is critical
2. profound understanding of the importance of school–community partnerships
3. a school culture built on high expectations for all students
4. coherent whole-school approaches to evidence-based literacy and numeracy teaching
5. building and sustaining teacher capacity to deliver whole-school practice
6. empowering, supporting and engaging Aboriginal and Torres Strait Islander students to enhance their learning capacity
7. making learning content engaging, accessible and culturally responsive.

Source: NCS (2012).

But there are major gaps in the education evidence base for Aboriginal and Torres Strait Islander students

A major limitation of the evaluation and research literature on education is that relatively few evaluations relate to programs designed specifically for Aboriginal and Torres Strait Islander students (and, in fact, much of the literature relates to overseas studies). Al-Yaman and Higgins (2011), for example, noted in one review of 139 evaluations relating to early childhood and schooling, that less than one in five of the programs and strategies were designed specifically for Aboriginal and Torres Strait Islander students.

Dockett et al. (2010) noted there was no solid evidence of the benefits of many early childhood interventions in Australia (and particularly of benefits relative to cost). The Closing the Gap Clearinghouse noted there are often small samples of Aboriginal and Torres Strait Islander people within evaluations of mainstream early childhood programs, and that many focused on output, rather than outcome, measures (AIHW and AIFS 2013).

There also continue to be gaps in what is known about the complex factors influencing school attendance and retention, with calls for more robust evaluations and better data,

including longitudinal data to track the progress of Aboriginal and Torres Strait Islander students (Bowes and Grace 2014; Helme and Lamb 2011; Purdie and Buckley 2010).

What is known is that attendance is a multi-faceted challenge that extends beyond truancy and absenteeism and involves issues of situational mobility and transience (and this requires targeted cross-sectoral responses). Attendance levels are also influenced by factors associated with the school environment, family attitudes to education, levels of engagement with students when they do attend and teacher attitudes (ACIL Allen Consulting 2014).

A recent report on *Strengthening Evaluation in Indigenous Higher Education Contexts in Australia* also commented that ‘there remains minimal publicly available evidence about program and policy effectiveness — that is, what does or does not work and why’ (Smith et al. 2018, p. 5). One of the report’s recommendations was that the Australian Government and philanthropic organisations mandate that a minimum of 10 per cent of all program funding in Indigenous higher education contexts be invested in evaluation.

Health

The evidence base for health policies and programs is stronger than other areas, although there are still significant knowledge gaps. As with education, there are a number of consistent themes coming through evaluations in the health area. These include:

- a holistic approach to health is required for Aboriginal and Torres Strait Islander people and communities (not just physical health of individuals but also the social, emotional and cultural aspects of communities)
- community participation in health programs is beneficial
- early screening for risk factors is likely to have net benefits, but there must be feedback provided regarding results
- taking a primary health care approach is likely to be beneficial at all stages of program development and delivery
- commitment is required by providers to developing culturally safe health promotion services
- there are benefits from developing collaborative partnerships between health agencies and others — for example, with outlets selling tobacco products and petrol
- supply reduction strategies (of potentially harmful products such as alcohol and drugs), demand reduction strategies and harm reduction strategies have all been successful
- policy makers need to take the history of colonialism into account in their responses to risky behaviours
- successful organisations employ Aboriginal and Torres Strait Islander staff who participate fully in program design, delivery and evaluation, and provide adequate training, where necessary, to strengthen capability

-
- there should be plans for built-in research, evaluation and monitoring to build an evidence base and provide a basis for advocacy for good practice interventions
 - there are benefits in encouraging youth services to support young people to engage in health promoting activities (AIHW and AIFS 2013; Al-Yaman and Higgins 2011; Bowes and Grace 2014; Closing the Gap Clearinghouse 2013; Mitchell, Bandara and Smith 2018; Osborne, Baum and Brown 2013).

There are gaps in the evidence base in the health area but the use of evidence is relatively strong

Al-Yaman and Higgins (2011) found a relatively high proportion of both Australian and Aboriginal and Torres Strait Islander research in the area of health. There was also a relatively high proportion of quantitative studies involving comparison groups, although relatively few randomised control trials. That said, the Closing the Gap Clearinghouse noted that ‘many evaluations of health programs did not use a comparison group, so it was difficult to draw definitive conclusions about the effectiveness of these programs’ (2013, p. 4). Bowes and Grace (2014) suggested the absence of comparison groups was due to inadequate funding for research, suggesting that without increased funding for research ‘much of the evidence base will continue to be of generally low quality’ (Bowes and Grace 2014, p. 3)

Al-Yaman and Higgins (2011) also noted many health programs shown to be effective in the mainstream population had not been evaluated for their impact on Aboriginal and Torres Strait Islander people. And a study by Eades et al. lamented the lack of meaningful research and evaluation looking at the actual health impact of health programs on Aboriginal and Torres Strait Islander people:

A more evidence-informed approach to improving Aboriginal health will also require research that goes beyond describing health problems to actually testing the impact of policies and programs. We located only 11 studies that tested approaches to improving health among urban Indigenous people, and a similar finding has been previously reported about Indigenous research generally in Australia and overseas. The paucity of information about the health of urban Indigenous people is conflated with the limited amount of intervention research in this area, resulting in a very small evidence base about what works to improve the health of Indigenous people. (2010, p. 522)

Eades et al. (2010) also noted that, although the majority of Aboriginal and Torres Strait Islander people lived in urban areas, only 11 per cent (their estimate) of Aboriginal and Torres Strait Islander health research included this group. Priest et al. (2009) and Bowes and Grace (2014) noted similar results for research into urban Aboriginal and Torres Strait Islander child health.

Issues with data availability also make evaluation of Aboriginal and Torres Strait Islander health programs more difficult. As Osborne, Baum and Brown stated:

Across all of the key determinant areas, there is a lack of high quality, publicly available evaluation data regarding programs and interventions, which limits the ability to identify success associated with such programs. (2013, p. 3)

Many of the problems in determining the success of health programs might be seen as inherent problems with evaluation. Many of the determinants of health and wellbeing are located outside the health system (for example, housing or employment), meaning that determining the impact of programs (relative to other factors) is very difficult. Further, people can also be affected by a number of government policies simultaneously and it is difficult to determine the relative impact of each program, particularly in areas affected by high levels of socioeconomic disadvantage (Muir and Dean 2017).

The problems can compound when undertaking evaluations. For example, a 2018 evaluation of a national anti-smoking campaign noted:

This level of long-term impact cannot be measured within the timeframe of this evaluation due to several factors including 1) the lack of baseline data, 2) the absence of population health surveys in the timeframe of the evaluation, 3) lack of time for population level behaviour change to occur and 4) difficulties in attribution. (Mitchell, Bandara and Smith 2018, pp. 10–11).

Housing

Poor quality or inadequate housing is associated with poor health and wellbeing outcomes, meaning it should be a focus for governments. Evaluations of housing programs have highlighted a number of elements of successful programs:

- addressing infrastructure, health promotion and the policy environment simultaneously
- enforcing appropriate standards and design guidelines while allowing sufficient flexibility to tailor designs and materials to local conditions
- ensuring high-quality, well-maintained health hardware such as taps, toilets, showers and sinks, coupled with attention to the safety of a house
- ensuring housing choices are responsive to community needs
- ensuring adequate food storage facilities and the availability of adequate potable water
- working collaboratively with Aboriginal and Torres Strait Islander communities and housing associations
- detailed planning of housing and housing-related infrastructure is critical
- training Aboriginal and Torres Strait Islander community members to gain skills and capacity
- ensuring cyclical maintenance of housing

-
- employing staff with cultural knowledge, and embedding best practice in cultural safety into policy settings and service delivery
 - ensuring the size and layout of houses meets the living needs of residents
 - improving indoor temperature regulation, as well as preventing damp, mould and fungi, reduces respiratory and skin diseases (Brackertz, Davidson and Wilkinson 2017; DPMC 2017b; Milligan et al. 2011; Osborne, Baum and Brown 2013; Ware 2013).

But housing related evaluations have their limitations

While housing programs have been extensively researched and evaluated, this cannot be said about the impact of housing programs on Aboriginal and Torres Strait Islander people. Even where this is more of a focus, evaluations usually have very small samples of Aboriginal and Torres Strait Islander people. The links between housing, health and wellbeing have also not been well researched. While housing projects can have diverse objectives — including improving health, increasing school attendance or reducing crime — these linkages are complex, often contentious, and programs are rarely evaluated against these broader objectives.

Ware identified there are clear associations between housing and health. Housing can affect health and wellbeing directly or indirectly through several pathways including physical, chemical, biological, economic and social factors. The impact of housing on health can be immediate (for example, electrocution, explosion, fire, structural collapse), or the effects may be delayed until later in life (for example, skin infections or eye diseases). Improved housing conditions can reduce health system costs, and improve participation in education and employment (Healthhabitat 2020; Ware 2013).

Recent research undertaken by the Australian Housing and Urban Research Institute for the Aboriginal Housing Office NSW concluded that:

There exists limited rigorous research that addresses the links between Aboriginal housing and non-shelter outcomes, with most evidence provided by the literature on housing and Aboriginal health outcomes. There is considerably less research on the links between housing and education, economic outcomes, social and community outcomes, empowerment and crime and safety. Consequently, where appropriate and available, findings from the literature on housing and non-shelter outcomes in the general population are included in this report as a proxy for research related specifically to Indigenous housing. (Brackertz, Davidson and Wilkinson 2017, p. 10)

And that:

The evidence base on effective programs to improve access to housing, housing security and stability for Indigenous Australians is relatively weak. This is partly due to the lack of rigorous program evaluations undertaken by jurisdictions and their limited public dissemination. Overall, the evidence base is spread thinly over a broad range of topics ... (Brackertz, Davidson and Wilkinson 2017, p. 10)

4 Types of evaluation

Key points

- Evaluations are initiated for a range of reasons, but the three key questions they seek to answer are: how well is the policy or program being delivered, what difference did it make, and do the benefits justify the costs? There are three types of evaluations to answer these questions:
 - process or formative evaluations — they assess whether a policy or program is being implemented the way it was intended and what can be learned and improved
 - impact evaluations — they test whether a policy or program has had an impact
 - economic evaluations — they compare the benefits of a policy or program with the costs.
- There is no one evaluation type, approach or method that can be applied to all circumstances. The evaluation type, approach and method adopted for policies and programs affecting Aboriginal and Torres Strait Islander people should depend on the questions that the evaluation is seeking to answer, the kind of policy or program being evaluated, the circumstances under which the policy or program is being implemented, and the resources available.
- More than one type of evaluation can be used, and in fact, different types can complement each other. For example, a formative or process evaluation undertaken in the early years of a policy or program (while the initiative works out start-up problems) could be followed by an impact evaluation after it is established (to judge its merit or worth).
- Employing triangulation (drawing from several sources of evidence) and using mixed methods (both qualitative and quantitative methods) can maximise the strengths, and compensate for the limitations, of any single method.

Policies and programs are put in place with the aim of improving outcomes for individuals and the community more generally. And while well-intentioned, policies often ‘rest on hypotheses rather than infallible laws’ (Marcel 2015). Hudson, in a report on mapping Aboriginal and Torres Strait Islander programs, said:

There is much goodwill in Australia to improve Indigenous outcomes. However, too many programs are implemented because of their perceived benefit, rather than a rigorous assessment of *a priori* evidence. (2016, p. 1)

Evaluation is about answering key questions about the effectiveness of policies and programs and identifying the factors that drive or undermine their effectiveness. It is more than the collection of data and information — it is a systematic assessment of the evidence which is then used to make a judgment about the design of a policy or program, its implementation and its outcomes.

Evaluation identifies hypotheses or theories of change and generates the data needed to test these. Done well, an evaluation will test the effectiveness of an initiative to deliver an intended outcome. As the Aotearoa New Zealand Evaluation Association & Superu put it:

At the end of an evaluation process, an evaluation needs to be able to say whether something is any good, or not, and why. (2015, p. 15)

The challenge is to evaluate in a way that produces credible, robust and useful results. Chapter 3 discussed some concerns about the rigour of policy and program evaluations. This chapter looks at the different types of evaluation, and the various approaches and methods that can be used when evaluating a policy or program (section 4.1). It also examines the factors to consider when deciding what type of evaluation to undertake, and issues raised by participants about the suitability of various approaches and methods for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people (section 4.2).

4.1 Evaluation types, approaches and methods

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

- 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?

Different evaluation types, approaches and methods can be used to evaluate policies and programs (box 4.1). The evaluation type and approach adopted will depend on the purpose of the evaluation, and the question(s) the evaluation is seeking to answer.

Box 4.1 **Key evaluation terms**

Evaluation type	The type of evaluation is defined by the evaluation question. The three main types of evaluation are process evaluations, impact evaluations and economic evaluations.
Evaluation approach	The approach used to answer evaluation questions. For example, an impact evaluation can use an experimental or quasi-experimental approach.
Evaluation method	Method refers to the way that information is collected and analysed to test theories and answer evaluation questions (for example, randomised controlled trials or propensity score matching).
Data collection	The collection of information to use in evaluation. This can be quantitative or qualitative.

Source: Adapted from HM Treasury (2020c, p. 7).

How well is the program or policy being delivered? Process evaluation

Process or *formative* evaluations are undertaken early in the development or implementation of a policy or program and seek to answer the question — how well is the policy or program being delivered? The purpose of these types of evaluation is to help form or shape a program to perform better. The evaluation assists in improving an initiative as it is rolled out and can provide the baseline for future evaluations.

Process evaluations usually involve program managers, evaluators, providers and participants working closely together, offering timely feedback to decision makers on how best to deliver a policy or program. Specific questions can include — is the program being implemented as intended? Is a policy appropriately targeted? How effective has implementation been so far? Can the policy or program be improved to achieve better outcomes?

Answering these questions allows an adaptive approach that supports learning by doing and adjusting policies and programs over time. Process evaluations typically collect quantitative data on inputs and outputs (and outcome data where feasible) and qualitative data from program recipients, providers and other stakeholders (collected, for example, via surveys or interviews). An example of a process evaluation for ‘1 Deadly Step’ is provided in box 4.2. This program was aimed at addressing the high prevalence of chronic disease among Aboriginal communities in New South Wales.

What difference did the policy or program make? Impact evaluation

Summative, *outcome* or *impact* evaluations typically make judgements about merit, worth or impact after a policy or program is well established. They seek to answer questions such as

— has the policy or program achieved its objectives? Is it effective? Has the policy improved outcomes? If so, by how much? Did the policy affect participants differently? Are there any adverse effects on those using the program? Impact evaluations are the most commonly recognised and an important form of evaluation because of their potential to influence policy and program decisions. Answering impact evaluation questions requires good data and statistical analysis to test what changes have occurred as a result of the policy or program.

Box 4.2 1 Deadly Step: A process evaluation of a chronic disease screening program

The 1 Deadly Step program^a aims to address the high prevalence of chronic disease among Aboriginal communities in New South Wales by increasing awareness and promoting prevention, early detection and timely health management. Developed by NSW Health in partnership with the Australian Rugby League, the program hosts public events where participants are taken through different stations to assess chronic disease risk factors, and uses a sporting platform to encourage participation. High-profile Aboriginal rugby league players from the local community are engaged as cultural ambassadors.

An evaluation in 2012 found high acceptability of the program among participating communities; however, improvements were recommended to screening and data collection processes and support for follow-up care. In 2015, the program was enhanced to include a web application for screening assessments, including a results portal for nominated care providers, a reporting portal for program administrators, and the ability to generate an evaluation report for each event site.

A process evaluation of the enhanced program adopted a three-stage approach.

1. A program logic was developed in consultation with key stakeholders, to identify core evaluation objectives and to design strategies and questions to measure objectives.
2. Data collection measures were implemented, and evaluators engaged with key stakeholders after each event.
3. Qualitative interviews were conducted, and quantitative clinical and survey data were analysed.

The program evaluation report contains quantitative data on the numbers and health profiles of communities participating in 1 Deadly Step. It also makes a range of recommendations to improve how the program is run, including event-related improvements and technical enhancements to the web application. The report recommends that evaluations in the longer term be conducted to examine the cost-effectiveness and impact of the program.

^a The term 'deadly' is used by many Aboriginal and Torres Strait Islander people to mean awesome, great, or excellent, and the name '1 Deadly Step' refers to participants making a step towards good health.

Source: Peiris et al. (2017).

The difficulty with assessing the impact of the policy or program is that outcomes can be influenced by factors not related to the initiative (known as extraneous variables). Most programs and policies will see improvements for some recipients in the outcome targeted by the program, but the improvements may not be because of the program in question. For example, a job training program aimed at helping people find employment may improve a recipient's job prospects, but so too could a person's prior work experience and general economic conditions.

One of the key challenges for impact evaluation is coming up with an estimate of the actual effects of a policy or program. To do this, evaluators need to estimate the effects of a program on policy and program participants (the 'treatment group'), and compare this with non-participants (the 'control group') who represent the counterfactual (box 4.3). Measuring the counterfactual requires:

- data of sufficient quality and quantity to support the analysis
- the counterfactual to be genuinely comparable to the intervention group
- that the effect of the policy or program is sufficiently large to be distinguished from 'noise' in the data (HM Treasury 2020c, p. 34).

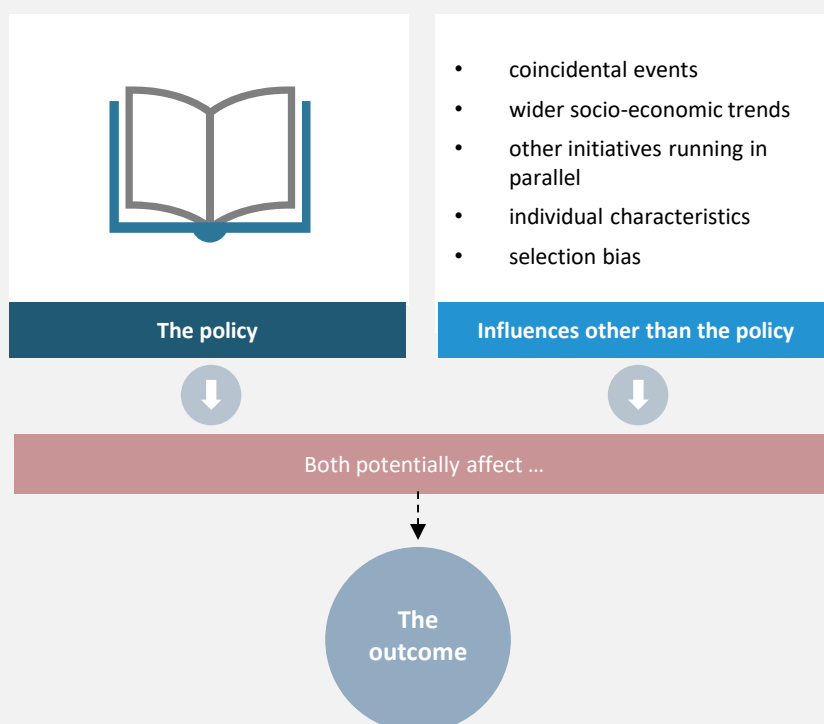
Experimental approaches, such as randomised controlled trials (RCTs), construct a control group through random assignment to assess the net impact of an initiative.²³ Usually, the average effect of the treatment on the treated population is compared with the effect on a suitably chosen control group, although more rigorous analysis attempts to understand why some treatments have an effect, on whom, and for how long. Sometimes an evaluation will capture variations in the intensity of services across the population, to examine the impacts of different service levels. An example of an RCT, which looked at the question of whether a web-based reading tool could increase literacy as effectively as regular instruction, is provided in box 4.4.

Evaluations using quasi-experimental and other quantitative designs also measure impact but are often considered weaker forms of evidence because of problems with internal validity and establishing the counterfactual (box 4.5).

²³ Some useful guides on experimental approaches to policy and program design and evaluation include: *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials* (Haynes et al. 2012); *Better Public Services Through Experimental Government* (Breckon and Mulgan 2015).

Box 4.3 What is influencing the outcome?

A key concept in impact evaluation is establishing the counterfactual — what would have happened if the policy or program had not been delivered, or does the outcome reflect the impact of the initiative in the absence of other influences?



The main strength of a randomised controlled trial (RCT) is that the control group serves as a point of comparison representing the counterfactual. There are several well-known threats to validity (extraneous variables) that can compromise the results of an evaluation, which RCTs seek to eliminate.

- *Differential selection*: where pre-existing differences between the treatment and control groups affect results (e.g. individual preferences, abilities, motivation, age or education level).
- *History*: where results are influenced by some other event that occurred during the delivery of the policy or program (e.g. a new company establishes itself, increasing local labour demand).
- *Maturation*: where results reflect naturally occurring changes because of the passage of time (e.g. participants become tired, older, or their skills improve).
- *Testing*: where the administration of a pre-test affects post-test results (i.e. participants become 'test wise').
- *Regression to the mean*: where the natural changes in the results of extreme cases are incorrectly attributed to the policy or program (e.g. an area with high crime rates records a lower crime rate for a period because of natural fluctuation).
- *Experimental mortality*: where differences in participants dropping out between the treatment and control groups influence results.
- *Causal order*: where results cannot establish whether an action preceded an outcome.

Sources: Farrington (2003, p. 53); HM Treasury (2011, p. 99); Mertens and Wilson (2012, pp. 308–309).

Box 4.4 Can a web-based learning tool increase literacy as effectively as regular instruction? A randomised controlled trial

Disparities in literacy skills between children can begin at an early age and have a profound impact on lifelong learning and educational attainment.

This randomised controlled trial looked at the question of whether a web-based reading tool could increase literacy as effectively as regular instruction with a randomised sample of 164 students in the Northern Territory. A control group of 148 students received literacy education through regular classroom instruction. Twenty-eight per cent of the total sample were Aboriginal or Torres Strait Islander students.

Students' literacy skills were pre-tested at the beginning of the semester, then students were randomly allocated to treatment and control groups, and tested again at the end of the semester. Treatment group students left their classes four times a week to attend 30–45 minute lessons with trained teachers using the literacy tool. The tool was designed to be especially engaging to children struggling to read, using rich animations and game-like learning activities.

The study found that all treatment group students made significant gains compared with their control group peers in two key reading competencies — phonological awareness and phoneme-grapheme knowledge. Similar improvements were observed among Aboriginal and Torres Strait Islander students in the treatment group, who appeared to accelerate in their early literacy skills to the point that they performed as well as their peers in phonological awareness despite lower overall attendance at school. However, because of the small sample size, the data lacked statistical power to establish the significance of differences between Aboriginal and Torres Strait Islander and non-Indigenous students. Aboriginal and Torres Strait Islander students, and students who spoke English as a second language, were also more likely to drop out of the trial.

Source: Wolgemuth et al. (2013).

Box 4.5 Quasi-experimental and quantitative designs

Quasi-experimental evaluations can involve statistical analysis using pre-test and post-test results, but they do not involve random allocation between control and treatment groups. They are simpler to administer than randomised controlled trials and are less likely to be limited by practical and feasibility constraints.

Type	Description	Pros and cons
Simple pre-post	Compares participants' outcomes before and after an initiative is implemented (i.e. pre-test/post-test, single group design).	<i>Pros:</i> Simple design. <i>Cons:</i> No counterfactual. Factors other than the policy or program can affect the result.
Simple difference	Compares outcomes of participants and non-participants after an initiative is implemented (i.e. post-test only, two group design).	<i>Pros:</i> Enables a comparison between groups. <i>Cons:</i> No randomised allocation. Pre-existing differences between groups can affect the result.
Simple matched sample	Compares outcomes of participants with those of a matched 'twin' with similar characteristics (allocation is not randomised).	<i>Pros:</i> Controls for certain known characteristics that can affect the result. <i>Cons:</i> No randomised allocation. Unknown factors can still affect the result.
Propensity score matching	Uses statistical probability analysis to allocate participants so that control and treatment groups have balanced characteristics (creates balanced samples in the absence of random allocation).	<i>Pros:</i> A method for forming balanced samples without randomised allocation. <i>Cons:</i> Unobserved, unmeasurable or unmeasured characteristics can still affect the result.
Regression discontinuity design	Individuals are ranked according to a given set of criteria, and a cut-off determines their participation in the policy. Participants just above and just below the cut-off are compared.	<i>Pros:</i> A method for forming balanced samples without randomised allocation. <i>Cons:</i> Findings may only hold for individuals around the cut-off. Participants could also change their behaviour around the cut-off (e.g. to gain access to the initiative).

(continued next page)

Box 4.5 (continued)

Type	Description	Pros and cons
Difference in differences	Compares participants' outcomes before and after an initiative is implemented with a relevant, non-participant population over the same period (but does not allocate participants to groups beforehand).	<i>Pros:</i> Enables a comparison between groups and eliminates some maturation effects. <i>Cons:</i> Pre-existing differences between groups can affect the result (and can change over time).
Multiple linear regression	Compares participants' and non-participants' outcomes using statistical operations, controlling for observable differences between the two groups that might affect their outcomes (e.g. gender, income or age).	<i>Pros:</i> Able to test and control for many variables without random allocation. <i>Cons:</i> Unknown factors can still affect the result.
Instrumental variables	Measures impact by identifying an 'instrumental' variable that affects the outcome of interest only indirectly by using it as the determining factor in whether an individual participates in the policy.	<i>Pros:</i> Uses rigorous statistical testing. <i>Cons:</i> Validity relies on finding a good instrument, or one that predicts the outcome only through participation in the initiative, which is difficult in practice.

Sources: Beal and Kupzyk (2014); OECD (2020a, p. 76).

Do the benefits justify the costs? Economic evaluation

Finding that a policy or program improves outcomes is not sufficient to determine that it is in the community's interest because taxpayer funds are limited and programs need to be judged against their costs. *Economic* evaluations look at the relationship between an initiative's costs and its effectiveness and seek to answer the questions — do the benefits of the program outweigh the costs, or was it worth it? Are there alternative approaches that would result in lower costs for the same benefits, or is the program good value for money?

Economic evaluation can be formative or summative, assisting decision makers to determine the most efficient policy solutions that offer the highest rates of return on investment (Crowley et al. 2018, p. 368). Specific methods include:

- *cost-effectiveness* analysis, which is a technique used to compare the relative costs to the outcomes of two or more courses of action. It is often used when evaluators cannot

conduct cost-benefit analysis because of constraints such as not being able to monetise benefits (Better Evaluation 2014). It involves quantifying the costs (in dollars) of delivering a policy relative to the quantity of outcomes achieved, producing an estimate of ‘cost per unit of outcome’ (HM Treasury 2011, p. 20). An adaptation of cost-effectiveness evaluation used in public health is cost-utility analysis, where outcomes are related to improvements in mortality and morbidity measured in Quality Adjusted Life Years (QALYs) or Health Years Equivalent (HYE) (Brouwer and Georgiou 2012, pp. 433–434).

- *cost-benefit* analysis is similar to cost-effectiveness analysis but attributes a monetary value to both inputs and outcomes so decision makers can assess whether the benefits outweigh the costs, as well as compare value across policies and programs (HM Treasury 2011, p. 20). Cost-benefit analysis can provide estimates of the change in value (positive or negative) derived from introducing a policy or program, calculated in terms of net present value or net social benefit (box 4.6).

The main strength of economic evaluation is that it uses a systematic and transparent framework for assessing the value of a policy or program. However, it can be difficult to value factors that are not easily quantified, and evaluators are often presented with incomplete or imperfect data from which to assess value. Other valuation techniques, such as Social Return on Investment (SROI), have also been developed specifically to take into account social and environmental value. Such methods emphasise a broad concept of value that measures inputs, outputs and activities from the perspective of the users of policies and programs. Engagement with stakeholders is a key part of the valuation process (Nicholls et al. 2012; Yates and Marra 2017).

Responses by Australian Government agencies to the Commission’s information request showed that few evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people include cost-benefit analysis or other economic analysis (chapter 3).

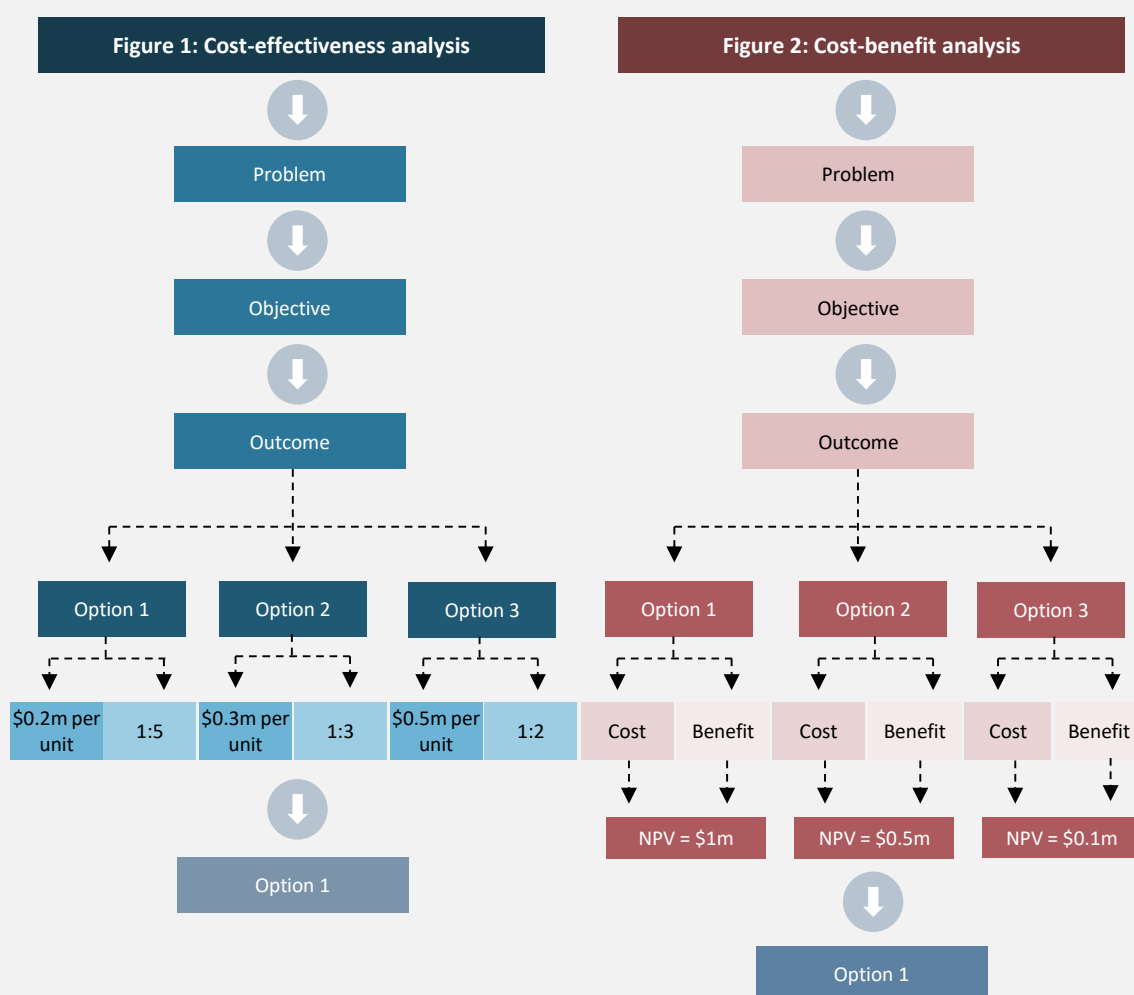
Aboriginal and Torres Strait Islander organisations are using economic evaluations to assess the value of initiatives. For example:

- the Healing Foundation commissioned a cost-benefit evaluation of the Murri School’s approach to delivering holistic education and wellbeing support to at-risk Aboriginal and Torres Strait Islander students by estimating the savings it achieves from diverting children from the child protection and youth detention systems, and the benefits from increased educational attainment and better mental health (Deloitte Access Economics 2017)
- the Maranguka Justice Reinvestment Project, which is a place-based, Aboriginal-led initiative, includes elements of cost-benefit analysis in its proposed evaluation approach (KPMG 2018; JustReinvest NSW, sub. 80). The initiative seeks to demonstrate the savings that could be achieved from directing funds away from crisis response programs, adult prison and youth detention, and reinvesting it toward education, support and prevention.

Box 4.6 Economic Evaluation Methods

Cost-effectiveness analysis (figure 1) compares the costs of different options relative to a specific outcome. Suppose a government has \$1 million and needs to decide how to spend it to maximise the benefit to the public. Option 1 would achieve the best outcome because, at an input cost of \$0.2 million per unit, the policy would obtain 5 units of that benefit.

Cost-benefit analysis (figure 2) calculates the monetary value of both inputs and outputs. The Net Present Value (NPV) of the investment estimates, in today's terms, the value that the investment will produce into the future. In this case, option 1 is the best choice because, for the same cost as the other options, it generates the highest NPV of \$1 million.



Source: Adapted from Department of Treasury and Finance (Vic.) (2013, pp. 43–44).

Various evaluation approaches

There are also new approaches to evaluation designed to support learning and adaptation and to use in complex systems, including developmental, realist, participatory and empowerment evaluation (box 4.7).

Box 4.7 **Evaluating in complexity**

Developmental evaluation

Developmental evaluation focuses on adaptive development — developing and improving an initiative because it is new or the context has changed, the clients of a program have changed significantly, or innovative solutions are needed to address persistent challenges. The evaluator is not an independent observer, but a core part of the program design team and involved in long-term progress and outcomes monitoring. Fundamental to the developmental approach are principles of *co-creation* (with policy and program managers and participants) and *timely feedback* (data collection, analysis and feedback is ongoing). The evaluator works in partnership with program managers and participants to:

- set up systems for evaluation, feedback and improvement
- collect and analyse data, and build stakeholders' capacity to use data
- understand system dynamics and uncover innovative strategies and ideas
- contribute evaluative data and logic to inform decision makers
- intervene to shape the course of the development of an initiative
- record lessons learned.

Realist evaluation

The general schema for realist evaluation testing is not cause and effect, but context–mechanism–outcome. The central task of evaluators is to understand what 'mechanisms' cause a desired outcome, and the contexts in which they work. In this way, the realist approach extends the basic evaluation question of 'what works' to 'what works, for whom, in what respects, to what extent, in what contexts, and how'. The realist evaluation process involves:

- developing an initial theory of change to describe how an initiative is expected to lead to an outcome, and under what conditions
- testing whether the theory of change is valid using relevant quantitative and qualitative methods
- looking for patterns of outcomes, identifying the mechanisms that generate those outcomes, and determining the contexts in which they occur (this can result in several context-mechanism-outcome hypotheses, applying to different situations or sub-populations)
- determining which hypotheses are most robust and plausible, and comparing these against the original theory of change
- adjusting the policy or program in line these new insights.

Sources: Patton (1994, 2016, pp. 526–527); Patton, McKegg and Wehipeihana (2015); Pawson and Tilley (2004); Westthorp (2014).

Developmental evaluation seeks to maximise the success of a policy or program through incremental review and improvement as it is rolled out, and relies on program participants' responses to shape learning and improvement. It is not concerned with program standardisation, or with an ex-ante and ex-post comparison of results, but instead 'aspires to continuous progress, ongoing adaptation and rapid responsiveness' (Patton 1994, p. 313).

This approach has grown out of observations about how innovation occurs in the corporate sector and the idea that answers are not known but can be discovered through a process of trying things (Gamble 2008).

Realist evaluation, a form of theory-based evaluation, seeks to understand the circumstances under which some programs are more or less effective with different groups of people. It seeks contextual explanations for why change occurs, while uncovering the diversity of participants' experiences. Box 4.8 provides an example of a realist evaluation of the Remote School Attendance Strategy.

Both developmental and realist approaches can be useful in 'complex situations', for example, where the relationships between inputs and effects are non-linear, where multiple interacting variables influence an outcome, and in cases of uncertainty and rapid change (Patton 2016; Rogers 2009a).

While these approaches can provide insights on how and why a program works and can identify areas for improvement, they do not measure causality or whether change has occurred compared with other scenarios. Findings are context-specific and are not generalisable.

Box 4.8 **A realist-informed and participatory approach to better understanding school attendance**

The Remote School Attendance Strategy (RSAS) is a place-based initiative focused on supporting school attendance. The Commonwealth and state governments deliver a range of programs under the initiative to support school engagement, which vary in each community. In 2017, Winangali (an Indigenous-owned engagement consultancy) evaluated the RSAS in partnership with commissioned evaluator, Ipsos. They chose a realist-informed, participatory and qualitative approach on the basis that this would be culturally appropriate, and would recognise context, complexity and diversity among families and communities. The evaluation involved:

- developing multiple theories of change on how families may respond to the RSAS, drawing from the literature, program data, anecdotal evidence and previous evaluations
- using local Aboriginal and Torres Strait Islander field researchers, and working with communities to identify a suitable researcher
- conducting 114 semi-structured interviews with parents and carers, and group discussions with parents, carers, local community members and RSAS staff
- researchers participating in a workshop to identify key themes, patterns and points of contention that arose during the fieldwork to develop and refine their understanding of how the initiative was working.

The evaluation found that the RSAS did not improve school attendance for all families and when acting in isolation it was unable to improve school attendance for families experiencing complex life events. However, the evaluation also found that the RSAS was an effective referral and coordination point, but only when staff were appropriately trained, knowledgeable about other services, and were seen as ‘the right person for the job’.

The evaluation identified four different family types served by the RSAS, each requiring different kinds of support:

- *committed families* — who see school as important and a stepping stone to higher education and jobs, but occasionally need practical support to get their children to school
- *protective families* — who want their children to attend school, but worry that school (or travel to school) is unsafe and so protect their children by keeping them at home
- *unsure families* — who are not convinced that education would lead to opportunities for their children, and worry that schools would not teach or value their culture
- *disconnected families* — who want their children to go to school, but are isolated and lack community ties, and do not know how to get children to school or who to ask for help.

The evaluation also found that different types of families responded to the same activity in different ways, and that a critical factor for the effectiveness of the RSAS was the ability of staff to recognise and respond to the different needs of families. As a result, program staff received additional training on the different family types and guidance on how to best support families.

Sources: ANAO (2019b); DPMC (2018d).

The evaluator's relationship with those involved with or interested in the policy or program can also be a defining factor in the evaluation approach.

- In an independent evaluation, the evaluator is primarily responsible for developing the evaluation question in collaboration with key stakeholders.
- With participatory or collaborative evaluation, stakeholders are integrally involved in the evaluation. The evaluator and the stakeholders jointly make decisions about the evaluation questions, how the evaluation is conducted, and how evaluation results are shared. 'Participatory evaluation' can occur at any stage of the evaluation process, can be practiced in various ways and can involve methods such as story-telling and participatory social mapping.
- With empowerment evaluation, the evaluator-stakeholder relationship is participatory and collaborative (box 4.9). The evaluator's role includes: consultation and facilitation directed toward democratic participation; and building the capacities of the participating stakeholders to conduct evaluations on their own, and to use the evaluation results (Rossi, Lipsey and Henry 2018).

Commenting on the more collaborative approaches, Rossi, Lipsey and Henry said:

Participatory evaluation, culturally responsive evaluation, and empowerment evaluation have principles that encourage respect and direct involvement of culturally diverse groups that themselves have been traditionally the subjects of evaluation. (2018, p. 294)

Box 4.9 **Empowerment evaluation: what is it?**

‘Empowerment evaluation’ involves a set of principles that guide the evaluation at every stage. The model aims to create a sense of ownership to ensure program outcomes will benefit participants. The approach is drawn from the participatory or collaborative field of evaluation and seeks to involve all stakeholders (evaluators, management, practitioners, participants and the community) in the evaluation process. Empowerment evaluation seeks to create positive change, as well as evaluating program outcomes — it can also ‘help people improve their programs and, in the process, their lives’ (Fetterman 2005, p. 43).

Empowerment evaluation situates control of the evaluation with the community. It is designed to foster a sense of ownership and to support organisations to be learning organisations. Fetterman argued that ‘ownership becomes stronger and more meaningful as a community uses its own evaluative findings to improve practice, finds evaluators supportive and helpful concerning the collection and interpretation of data, and experiences the trust of funders to take additional time (and the requisite detours, within reason) to accomplish its objectives’ (2005, p. 44).

Capacity building is one of the features of empowerment evaluation. Accordingly, empowerment evaluation is said to be ‘particularly suited to social justice programs working with oppressed people, especially where the aim is to increase self-determination and create positive change [including] ... programs that work with Indigenous peoples, culturally and linguistically diverse groups, and people who have experienced family violence or sexual assault’ (Australian Institute of Family Studies 2015).

One of the criticisms of empowerment evaluation is that the process is resource and time intensive for organisations. Another is that empowerment evaluation can be more a method of teaching people about evaluation than an evaluation itself.

Sources: AIFS (2015); Fetterman (2005).

Data collection methods

Policy makers are also interested in ‘why’ questions — why was implementation more or less successful in certain cases, why did the policy not make the difference policy makers thought it would, why were the costs so much higher (or lower) than expected, or why did the policy benefit some people, but not others?

Qualitative methods, such as interviews, focus groups and case studies, can provide in-depth insights into people’s experiences and perspectives (box 4.10). They have a particular strength in addressing questions of why and how (WHO 2013, p. 34).

Citing a literature review on culturally competent evaluation by Chouinard and Cousins (2007), Muir and Dean noted that ‘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 rather than the imposed categories of a survey or psychometric measure’ (2017, p. 8). And as Liamputtong said, ‘qualitative methods can help counteract power differences by giving voice to those with less power — to hear people’s voices, to share their stories, [and] to empower them’ (2013, p. xii).

Box 4.10 Strengths and weaknesses of various qualitative methods

Qualitative methods can provide additional information to quantitative methods. They often require direct involvement with policy or program participants and/or those involved in administering the initiative. Evaluators engage in detailed conversations or work closely with others to elicit information from the perspectives of individuals, groups or whole communities.

Type	Description	Pros and cons
In-depth interviews	Information is collected through detailed questioning and conversation with an individual.	<i>Pros:</i> Provides detailed information from the individual's perspective and allows the evaluator to probe and gather additional information. <i>Cons:</i> High cost, time consuming, and relies on strong interviewing skills. Also risks biased results (some individuals are more likely to participate in interviews).
Case studies	Describes 'real life' impact by exploring a single case in depth. Subjects are often selected because they are unusual rather than being representative. Case studies can use multiple sources of evidence.	<i>Pros:</i> Able to explore real life situations in depth. Can help understand complex phenomena. <i>Cons:</i> It is difficult to generalise findings to national/population level or to different contexts, situations or phenomena.
Focus groups	Involves an in-depth discussion with a small group of people about a specific topic. Group interactions are a critical part of information gathering.	<i>Pros:</i> Less costly than individual interviews but can be time intensive. <i>Cons:</i> Risk of bias (some individuals are more likely to participate in focus groups). Can create group think. Less able to explore complex matters.
Narrative enquiry	Involves collecting 'narratives' (people's stories and oral histories) to understand the meaning of individual and collective experiences.	<i>Pros:</i> Provides rich information on participants' experiences. <i>Cons:</i> High cost and time consuming. Requires substantial commitment from participants.
Ethnography	Draws information from peoples' lived experiences, daily activities and social contexts. Requires close involvement with participants for a lengthy period.	<i>Pros:</i> Enables an in-depth examination of participants' world views. <i>Cons:</i> Imposes heavily on participants' lives. Requires long timeframes.

Sources: HM Treasury (2020c, pp. 42–43); Liamputtong (2013); Mertens & Wilson (2012, pp. 330–349).

The Tasmanian Government commented that it:

... has found that qualitative approaches to Aboriginal and Torres Strait Islander-specific policy and program evaluation are useful in compensating for lack of high-quality, accessible quantitative data, with the added benefit of providing for a deeper understanding of Aboriginal and Torres Strait Islander people's unique experiences. (sub. 100, p. 4)

Similarly, Colleen Hayward said:

Quantitative [data], as you know, is just going to give you the numbers. They don't tell you the story behind it and in my view, it is the stories behind the numbers where the richness actually lies ... (in Smith et al. (2018, p. 36))

Quantitative methods (such as surveys and questionnaires) can collect a lot of information in a short period of time, although questions should be pilot tested to ensure they are not misunderstood by respondents. As Muir and Dean said:

... care has to be taken to ensure that people understand the questions they are being asked. Survey questions, measurement tools and psychometric tests that were developed for largely urban and non-Indigenous populations might not be appropriate for people with different world views. (2017, p. 8)

Another issue for quantitative methods is that in small communities there may not be enough program participants for quantitative methods to produce statistically meaningful results (sampling is discussed further in chapters 5 and 9).

Triangulation across sources of data can make each data source more valuable and strengthens validity (box 4.11). Examples of data triangulation include:

- surveys that ask a mix of scale questions (quantitative) and open-ended questions (qualitative)
- assessment tools that include free-text notes as well as standardised quantitative measures
- asking different people about the same issues, for example, getting feedback from students as well as teachers
- corroborating self-reported data with independent observation, for example, about students' mastery of a new skill or teachers' confidence with a new teaching practice (Department of Education (NSW) 2019).

Box 4.11 About triangulation

The term triangulation comes from orienteering — to locate yourself on a map you need to take your bearing from at least two reference points. The strongest geometric shape is also the triangle (Patton 1994).

Denzin (1989) identified four types of triangulation.

- Data triangulation — combining data from more than one source, for example, from several settings, points in time or groups of people.
- Methodological triangulation — combining different research methods. This can include ‘within research’ triangulation (where a range of different lines of questioning might be used to approach the same issue) and ‘between method’ triangulation (where different data collection methods are combined). This can also include the combining of qualitative and quantitative data.
- Investigator triangulation — this involves more than one researcher looking at the data so that they can either check or challenge each other’s interpretation or deliberately approach the data from different angles.
- Theory triangulation — this means looking at the data from different theoretical positions in order to explore the fit of different theories to the data, and to understand how looking at the data from different assumptions affects how it is interpreted.

Sources: Denzin (1989); Patton (1994).

4.2 How do you decide on an evaluation type, approach and method?

The evaluation type and approach used should depend on the evaluation questions, the kind of policy or program being evaluated, the context or circumstances under which the policy or program is implemented, and the time and resources allocated to the evaluation.

Questions to ask when deciding on a choice of evaluation include:

- 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 control is important 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 intervention being implemented (will it allow the estimation of a counterfactual)? (HM Treasury 2011, p. 24).

More than one type of evaluation can be used, and in fact, different types of evaluation can complement each other (the evaluation of the Get Healthy Information and Coaching Service is an example — box 4.12).

A formative or process evaluation undertaken in the early years of a policy or program (while a new policy or program works out start-up problems) could be followed by an impact evaluation after the program is up and running (to judge the merit or worth of the program). Similarly, an evaluation could combine an experimental approach to assess impact with a comparative case study to better understand different contexts (Stern 2015). It is rare for evaluations to ask only a single question. It will also be rare for a single approach to be able to answer all the evaluation questions. Evaluations assessing impact often also want to know what works, when and in what circumstances.

Both quantitative and qualitative data have a place in evaluation and together can tell a more complete story (chapter 9). Quantitative data can be used to track trends over time and understand patterns and differences in outcomes between policy and program participants. As noted earlier, qualitative data can help provide context and answer ‘why’ questions.

Participants to this project were supportive of mixed approaches and methods for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people on the basis that they provide both empirical and contextual insights on the effects of policies and programs (box 4.13).

The time and resources available for evaluation are also important considerations when choosing an evaluation approach and method. RCTs and quasi-experiments, for example, can be expensive to run.

Box 4.12 Evaluation of the Get Healthy Information and Coaching Service Aboriginal program

The Get Healthy Information and Coaching Service (GHS) provides free telephone-based support to help adults at risk of chronic disease to improve healthy eating, increase physical activity, reduce alcohol consumption, and achieve a healthy weight. The NSW Ministry of Health (MoH) sought to introduce a dedicated Aboriginal GHS, and to inform its development, initiated a three-stage, multi-method evaluation (completed over five years).

1. Formative evaluation

A formative evaluation was conducted to understand how the mainstream GHS could better meet the health and cultural needs of Aboriginal people. A Working Group comprising the Aboriginal Health and Medical Research Council, the NSW Ministry of Health, and medical and health professionals guided the formative evaluation. Evaluation methods included phone and face-to-face interviews with participants (subject to voluntary consent), group discussions with Aboriginal people who were no longer participating, and interviews with Aboriginal community leaders, health professionals and service providers. The evaluation found that the service was beneficial but should cater to the specific needs of Aboriginal people — weight loss and consideration of food choices were not high priorities in Aboriginal communities. Findings suggested that promotional materials should emphasise that GHS is a free and confidential service, and that the MoH should work closely with Aboriginal Community Controlled Health Services (ACCHSs) to deliver the program. The MoH made changes to the proposed service model in light of this evidence, including implementing a participant recruitment strategy involving promotions and referrals through ACCHSs.

2. Appropriateness analysis

Appropriateness analysis assessed whether GHS was suitable for Aboriginal participants, including cultural understanding and respect. The MoH commissioned the Cultural and Indigenous Research Centre Australia (CIRCA) to conduct the appropriateness analysis, with ethics approval obtained from the Aboriginal Health and Medical Research Council Ethics Committee. Potential participants (n=101) were sent a letter and information sheet and then contacted by phone to seek their approval and informed consent to participate in the study. Interviews were conducted by CIRCA research consultants (including Aboriginal research consultants). CIRCA used thematic analysis to identify key themes and assess the appropriateness of proposed enhancements to the mainstream GHS for Aboriginal people.

3. Quantitative impact analysis

Quantitative analysis measured the impact of GHS on participants' health and physical activity. The MoH conducted a pre-test and post-test evaluation using participant data collected at enrolment and completion of the program. Measures included self-reported weight and height (to calculate a body mass index), waist circumference, and levels of physical activity. Overall, participants who completed the targeted 6-month Aboriginal GHS made significant improvements on average against health measures, with no significant differences compared with non-Aboriginal participants of the mainstream program. However, participants in the Aboriginal program were significantly more likely to be from the lowest socio-economic quintile, and live in regional, remote or very remote locations, compared with participants in the mainstream program. This suggested that the targeted service was increasing access to Aboriginal people who experienced greater disadvantage. Overall, the proportion of Aboriginal participants in GHS increased from 2.3 per cent (n=66) in 2009 to 8.8 per cent (n=345) in 2015.

Sources: O'Hara et al. (2016); Quinn et al. (2017).

Box 4.13 **Support for mixed approaches and methods**

NSW Aboriginal Land Council

NSWALC is of the view that no one method (qualitative or quantitative) should be favoured over another as each method elicits different information which all help to aid understanding. Quantitative data is helpful for a bird's eye view and to see patterns, while qualitative data is vital to provide context and to explain the why or how. Neither should be viewed as inferior or superior to the other. (sub. 91, p. 5)

Queensland Aboriginal and Islander Health Council

Evaluation methodologies should recognise a variety of methods that are context-specific ... it is unrealistic to expect randomised controlled trial (RCT) methodologies if this evaluation method is not acceptable or feasible ... The criteria for feasibility is also context-specific, in that Aboriginal and Torres Strait Islander stakeholder priorities should guide the optimal selection of methods, just as economic considerations do. (sub. 55, p. 8)

The Western Australian Government:

A mixed evaluation approach including both quantitative and qualitative methods may provide the best insights. (sub. 74, p. 2)

National Indigenous Australians Agency

NIAA encourages a broad based approach to undertaking evaluations, which draws upon a range of evaluation traditions. Under the [Indigenous Advancement Strategy] Evaluation Framework, NIAA bases the choice of evaluation methods on a range of factors including the specific evaluation question(s) and body of existing evidence ... and the need to integrate Indigenous community values, knowledge and perspectives to ensure findings are useful, credible and helpful ... We also promote the use of mixed methods to ensure adequate coverage of Indigenous perspectives and to generate findings relevant and useful to both Indigenous communities and government. (sub. 86, p. 7)

Tony Dreise et al.:

Having clear theoretical and epistemological foundations for mixed methods evaluation supports more comprehensive and robust forms of analysis and strengthens the credibility and relevance of the evaluation findings. ... We believe that rigorous evaluation requires the integration of Indigenous and Western methods of knowledge production. (sub. 33, p. 11)

Are some types and approaches more or less suited?

Some participants suggested that certain approaches to impact evaluation, including experimental approaches such as RCTs, were less suited to understanding policy and program outcomes for Aboriginal and Torres Strait Islander communities. For example, the Western Australian Government said:

While [RCTs] are considered the 'gold standard' in evaluation methodology, there are practical considerations and sensitivities associated with creating 'treatment' and 'control' groups in the context of policies and programs impacting Aboriginal and Torres Strait Islander communities. RCTs also tend to be very expensive and resource intensive. (sub. 74, p. 2)

Tony Dreise et al. noted:

... while [RCTs] are viewed as a gold standard for causal inference, they are not appropriate in every context ... we stress that a broader range of variables be considered in order to capture Indigenous rights, perspectives and aspirations within evaluation. (sub. 33, p. 8)

And the National Aboriginal Community Controlled Health Organisation said:

Less appropriate approaches include [RCTs] and other experimental designs due to ethical dilemmas of one cohort accessing a service and another cohort not — when all participants require the service. (sub. 95, p. 3)

RCTs are used widely in areas of medical, health and other scientific research, often where treatments are homogenous and defined, and where the nature of experiments are largely under the control of researchers. They are used less in evaluation for social policy where programs are designed to deal with multiple and complex issues and where change is expected to occur in the long term, although their use is increasing (Ames and Wilson 2016; Haynes et al. 2012). Figure 4.1 outlines the conditions under which experimental approaches, such as RCTs, are more and less suited.

Many of the ‘less suitable’ conditions for experimental approaches will apply to evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people and communities. This includes, for example, where there are small sample sizes, where the effects of a policy or program are difficult to disentangle from the effects of other programs, where a complex relationship exists between the activities of a policy or program and its expected outcomes, or where it is not practical or ethical to create different comparison and target groups.

On the issue of small sample sizes, the Commission’s study on *Expenditure on Children in the Northern Territory* noted that randomisation is challenging in small communities because of biases that may be introduced as a result of:

- peer effects: people from the control group are aware that they have been omitted from the program, and the mere existence of the program may alter their behaviour. And people from the treatment group may share learnings with others in the community, which would reduce the measured difference between the groups
- selection bias: people experiencing significant challenges and disengagement from support (who may be the target of a service) may be less willing to participate in the trial (if participation is voluntary), which would reduce the measured size of the effect. In contrast, people who are more motivated and predisposed to success may persuade the program provider to allow them to participate, leading to an overestimate of the program’s effectiveness (2020, pp. 277–278).

Figure 4.1 When to use quantitative impact evaluation

		More suited if...	Less suited if...
Nature of the policy or program	➔	The initiative is new, or a distinctive change is made to an existing one Activities are standardised for all participants The initiative and its objectives are distinctive	Changes to the initiative are small or incremental Activities vary between participants Many overlapping initiatives contribute to similar objectives
Type of delivery	➔	The initiative can be delivered selectively, or phased in, so that participants can be allocated to treatment and control groups	Selective delivery is undesirable or not possible Delivery is universal or launched simultaneously nationwide
Expected impacts	➔	A direct, causal relationship is expected between an initiative's activities and outcomes Effects are immediate or expected in the near-term The expected impact of the initiative is large compared to other influences The influence of extraneous factors can be controlled for	Activities and outcomes are only weakly or indirectly linked Effects are gradual, achieved in the longer term The expected impact of the initiative is small compared to other influences Many complex, interrelated factors influence outcomes that cannot be controlled
Data quality	➔	Baseline data are collected prior to roll out Unit-level data are collected at precise time points Variables are easily quantified	Data are collected retrospectively Data are aggregated and collected at imprecise times Variables are difficult to quantify
Population of interest	➔	A large population exists from which to draw samples	Working with small populations, sub-populations or participant groups

Sources: Ames and Wilson (2016); Rogers (2009b); White, Sabarwal and de Hoop (2014). Adapted from Campbell and Harper (2012, p. 7).

Also, because RCTs can hide asymmetric effects across a study population, especially in heterogeneous samples (Deaton and Cartwright 2018), evaluators need to consider carefully how and when to use RCTs to evaluate policies and programs affecting Aboriginal and Torres Strait Islander people.

That said, there will be circumstances where RCTs are suitable — as evidenced by a number of RCTs for policies and programs affecting Aboriginal and Torres Strait Islander people (see: box 4.4; Goldstein and Hiscox (2018); Ward et al. (2013)). For example, by leveraging the inability of a school attendance program to serve every eligible student, Goldstein and Hiscox (2018) were able to compare the outcomes of eligible students (who were assigned to both control and treatment groups) in their RCT of the Australian Government’s School Enrolment and Attendance Measure.

In terms of design rigour, Deaton and Cartwright (2018) cautioned against a simplistic reliance on RCTs and suggested that alternative methods also yield strong evidence. They argued that random assignment does not automatically provide unbiased samples, and that explicit controls of known extraneous effects is favourable over randomisation to remove bias (for example, by collecting matched samples so that treatment and control groups have similar characteristics, and using statistical validation methods to test for bias).

As with any experiment involving people, ethical principles relating to respect and protection from harm are foremost concerns. In many cases, ethical issues can be addressed through careful formulation of an appropriate evaluation question and RCT design (box 4.14). For example, ethical issues can arise from withholding the potential benefit of an initiative from control groups in an RCT. A ‘stepped wedge’ design can be used to overcome this issue. This is where everybody eventually participates in the program through phased allocation to the treatment group, but the order in which participants are allocated is determined randomly.

Box 4.14 Randomised controlled trials and ethical issues

As with any research involving people, ethical principles relating to respect and protection from harm are important concerns. Specific ethical issues arising from randomised controlled trials (RCTs) relate to:

- the lack of benefit for the control group
- the potential for harm to either the control or treatment group, since the outcomes of involvement in a policy or program (or abstaining from it) are not known
- problems with informed consent, particularly if providing information to participants could bias the results. Individual consent may be meaningless where the nature of the policy or program itself is such that it is impossible to exclude individuals from an RCT (for example, for some compulsory or compliance-oriented programs).

Such ethical problems do not rule out using RCTs as a legitimate policy evaluation tool. In many cases, ethical issues can be addressed through careful formulation of an appropriate evaluation question and RCT design. Some ethical concerns may also be addressed by:

- forming control and treatment groups by staggering the roll out of an initiative (stepped-wedge design) or making use of waiting lists in the presence of program capacity constraints
- explaining clearly to participants the purpose and benefits of randomisation (to the extent possible), and compensating control group participants if their active involvement is needed for data collection
- stopping an RCT if it becomes clear that harm is occurring or if treatment and control groups are achieving substantially unequal outcomes
- offering alternative assistance should the experimental treatment fail.

Sources: Haynes et al. (2012); Osrin et al. (2009); Posavac and Carey (1997); Weijer et al. (2011); White, Sabarwal and de Hoop (2011).

Other evaluation designs can have wider application when random allocation of program participants to control and experimental groups is not possible for practical reasons (box 4.5 outlines the strengths and weaknesses of the different approaches). Tony Dreise et al., for example, suggested that statistical designs such as difference-in-difference, instrumental variables estimation, propensity score matching and regression discontinuity are approaches that could provide results comparable to RCTs if data are collected in a suitable form (sub. 33, p. 8).

Are other evaluation approaches better suited?

Developmental, realist and participatory evaluation were highlighted by participants as being particularly suited to evaluating policies and programs affecting Aboriginal and Torres Strait Islander people (box 4.15) on the basis that they:

- are better suited to complex problems
- can be more effective at incorporating Aboriginal and Torres Strait Islander knowledges, perspectives and world views into evaluations

-
- can engage Aboriginal and Torres Strait Islander people throughout the evaluation process, including using culturally sensitive methods, such as yarning (storytelling), ganma (knowledge sharing) and dadirri (listening, box 4.16)
 - can build participants' evaluation capabilities.

As discussed earlier, realist evaluation seeks contextual explanations for why change occurs, while seeking out diversity of participants' experiences.

The exploratory nature of this approach supports participatory processes, as well as learning and improvement to policy and program design. It can be useful when:

- the outcomes of a new initiative, pilot program or trial 'that seem to work but for whom and how is not yet understood'
- existing data or previous evaluations deliver mixed results (to understand how or why this has occurred)
- evaluating programs that will be scaled out (to understand how to adapt the initiative to new contexts) (Westthorp 2014, p. 7).

Developmental evaluation seeks to maximise the success of a policy or program through incremental review and improvement as it is rolled out (relying heavily on program participants' responses to shape learning and improvement). This 'learning by doing' approach may be particularly useful in remote Aboriginal and Torres Strait Islander communities, where both local context and supporting capability building are important. Moran, whose field of expertise is community development, commented that:

There needs to be a recognition that conventional evaluation frameworks do not actually tell you what happened ... This is not very helpful to community development, where people need help to incrementally step their way through complex problems, and to be always looking to what is the best next step. Frontline workers need to be able to make small evaluations over short periods of time, to consider what the results were from their last steps. Evaluation efforts then focus on the next step, whether a new tactic is required, or if this method should be continued. (Moran 2017, p. 3)

Box 4.15 **Support for developmental, realist and participatory evaluations**

Darling Downs and West Moreton Primary Health Network:

The Realist evaluation approach when the participants or community have been involved in the design and implementation of the program will be most successful. It will have the strongest sense of ownership and empowerment. The other approaches are important such as ... developmental evaluation, participatory evaluation and economic evaluation ... and should be included into a large whole of program review. (sub. 6, pp. 1–2)

Australian Institute of Aboriginal and Torres Strait Islander Studies:

We concluded in a 2017 literature review, that *qualitative* and *participatory approaches* are often most appropriate for work in Indigenous research, as it is most consistent with the demands of decolonisation, including recognising and foregrounding Indigenous epistemologies and perspectives. Participatory approaches assist in avoiding some of the problems of other evaluative models which require the imposition of top-down, pre-determined categories that typically, and historically, have failed to capture reality for many Aboriginal and Torres Strait Islander people. (sub. 72, p. 12)

Australian Evaluation Society:

Some evaluation approaches that enable ... self-determination and agency may include case based, participatory, collaborative approaches, e.g. empowerment, democratic, strengths based, place-based evaluation approaches. Again, a partnership approach with Aboriginal community representatives would encourage self-determination and valuing of Aboriginal and Torres Strait Islander knowledges, as part of the processes adopted in an evaluation. (sub. 49, p. 18)

The Western Australian Government:

Approaches such as developmental evaluation, participatory evaluation and realist evaluation, can sit well in evaluating programs affecting Aboriginal and Torres Strait Islander people. Some approaches incorporate methods that are culturally sensitive, such as storytelling and spending time with Indigenous people to gain a deeper understanding of issues. The choice of evaluation approach should be made on a fit-for-purpose basis in each instance. (sub. 74, p. 2)

Victorian Aboriginal Child Care Agency:

In terms of broad approaches to evaluation, VACCA supports those that are transformative and participatory. In terms of specific methods, VACCA believes that several qualitative methodologies, such as Most Significant Change, narrative, action research and case studies can serve as culturally appropriate and meaningful. VACCA is supportive of mixed methods approaches that may be inclusive of Western methods, provided these are endorsed in a highly collaborative design process and accompanied by other culturally meaningful methods. (sub. 26, p. 4)

Department of the Environment and Energy:

... 'Participatory' and 'Realist' evaluation approaches ... are likely to be the most suited to policies and programs affecting Indigenous people, as they ensure stakeholders and program participants are involved in the evaluation process, and seek to understand *how* an intervention leads to change. (sub. 101, p. 3)

Tony Dreise et al.:

If conducted effectively and ethically, participatory evaluations will enable Indigenous knowledge, perspectives and world views to be incorporated into program and policy evaluation. Furthermore, community-based and participatory evaluation methods have 'spin-off' benefits beyond improving the programs and policies under evaluation; they also build community capability in the research field and beyond. (sub. 33, p. 4)

Box 4.16 **Yarning and dadirri**

Narratives and stories are a fundamental part of sharing Aboriginal and Torres Strait Islander ways of knowing.

Yarning is a qualitative research method that draws on Aboriginal and Torres Strait Islander people's cultural practice of storytelling as a way of conveying information and for relationship building (it can be one-on-one or in groups). It involves in-depth discussions to gain 'thick description' and offers a culturally safe place for Aboriginal and Torres Strait Islander people to freely talk about their experiences. Yarning positions the evaluator as a listener and learner in the data collection process and respects Aboriginal and Torres Strait Islander people as the authority of their knowledges. It involves reciprocal relationships and is a two-way process of learning and knowledge exchange.

The Victorian Aboriginal Community Controlled Health Organisation described the different types of yarning:

... 'social yarning' (which builds trust and engagement), 'research topic yarning' (a relaxed interview that focuses on information gathering), 'collaborative yarning' (sharing ideas or brainstorming about the research topic), and 'therapeutic yarning' (participant disclosure of emotional or traumatic experiences). The diversity of types of yarning allows it to be applied to various elements within the process of evaluation. (sub. 44, p. 15)

Dadirri is an Aboriginal concept and practice of deep contemplation and listening. The word comes from the Ngan'gikurunggurr and Ngen'giwumirri languages (meaning 'listening'). Dadirri has also been described as both an 'indigenous way of life' and 'a research methodology'; and as a methodological approach, indigenous philosophy informs the investigative process, and ensures ethical behaviour and cultural safety in research design. It is an approach that:

- honours the integrity and fidelity of Aboriginal communities
- requires reciprocity in how researchers relate to, and behave with, Aboriginal people
- encourages researchers to reflect on their own assumptions or bias
- uses Aboriginal world views in which learning is a process of listening and witnessing without judgement or prejudice, and of being responsible for self in relationship with others in the listening and learning process
- upholds Aboriginal world views so that the activity of learning introduces a responsibility to act with integrity and fidelity to what has been learnt (Atkinson 2002, p. 20).

Sources: Atkinson (2002, pp. 15, 20); Geia, Hayes and Usher (2013, p. 16); Leeson, Smith and Rynne (2016, p. 8).

Indigenous approaches

Indigenous evaluation approaches place an Indigenous lens over the initiative being evaluated, as well as the evaluation process itself, to ensure that the values, experiences, knowledges and interests of Aboriginal and Torres Strait Islander people are recognised and upheld. There is no single, definitive Aboriginal and/or Torres Strait Islander approach, but Aboriginal and Torres Strait Islander people, academics and organisations have developed frameworks and methods for evaluating policies and programs. These approaches are

generally principles-based, promoting Aboriginal and Torres Strait Islander people's self-determination as well as strengths-based, culturally-relevant and respectful processes.

One example is Ngaa-bi-nya — a program evaluation framework developed by Aboriginal academic and health researcher, Megan Williams, in collaboration with Aboriginal and Torres Strait Islander program providers and staff (Williams 2018). Ngaa-bi-nya provides a structure to analyse a program's capacity to reflect Aboriginal and Torres Strait Islander people's preferences and contexts across four domains — landscape, resources, ways of working and learnings (box 4.17).²⁴

In New Zealand, there is a growing emergence and use of Māori evaluation theories, such as Kaupapa Māori (which means 'Māori way', see Smith 1999; Wehipeihana 2019a). Kaupapa Māori evaluation can be characterised broadly as evaluation that is:

- controlled and owned by Māori
- conducted for the benefit of Māori (although it may also benefit others)
- carried out within a Māori world view, which is likely to question the dominant culture and norms
- aims to make a positive difference for Māori (Smith (1999); and Moewaka Barnes (2000), in Moewaka Barnes (2013, p. 165)).

Cram, Pipi and Paipa describe Kaupapa Māori evaluation as being about:

... the reclaiming of knowledge creation mechanisms. This can be done by growing Māori-centred explorations for Māori marginalization; building knowledge about Māori practice models; and developing theories about the reasons, directions and timeliness of positive changes that occur in the lives of Māori because of interventions. (2018, p. 68)

A number of Aboriginal and Torres Strait Islander scholars — such as Rigney (2006), Martin (2003), Moreton-Robinson (2004), and Nakata (2002) — have challenged Western research approaches on the basis that they fail to recognise the knowledges, views and interests of Aboriginal and Torres Strait Islander people, and have encouraged Aboriginal and Torres Strait Islander people to develop new research methods that are culturally relevant and that give Aboriginal and Torres Strait Islander people voice. For example, Rigney, noting that there was not a unanimous view on method, methodology and epistemes in conducting research with Indigenous peoples, said that:

... what is central to Indigenist Research is that Indigenous Australian ideals, values and philosophies are core to the research agenda ... Indigenist researchers are united in the belief that it is inappropriate that research contribution to Indigenous social movements should come solely from non-Indigenous Australians. (2006, p. 41)

²⁴ Another example is *An Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health* (Kelaher et al. 2018a), and the associated *User's Guide* (Kelaher et al. 2018b), by the Lowitja Institute and University of Melbourne.

Box 4.17 Ngaa-bi-nya program evaluation framework

Ngaa-bi-nya is a program evaluation framework that provides a structure to analyse a program's capacity to specifically reflect Aboriginal and Torres Strait Islander people's preferences and contexts by providing prompts for evaluators to consider across four domains.

- *Landscape factors* examine broader systems that influence the capacity of programs to deliver care. This includes looking at the history, environment, services and other aspects of the local area and experiences of the traditional owners, and the differences between the lives of Aboriginal and Torres Strait Islander people compared with others in that place.
- *Resources* include the physical, financial, human, and material resources that often support Aboriginal and Torres Strait Islander programs. Evaluation looks at whether resources in each area are adequate to meet the needs of Aboriginal and Torres Strait Islander people.
- *Ways of working* refers to the extent to which culturally relevant data collection tools are used, and the involvement of Aboriginal and Torres Strait Islander people in leading evaluation design, data collection, analysis, reporting, and knowledge exchange.
- *Learnings* prompts evaluators to consider insights gained and lessons learnt. It includes thinking about the extent to which the program contributes to: an evidence base; self-determination; strengthening and preserving culture; and opportunities for building the capacity of Aboriginal and Torres Strait Islander staff and community members in research, monitoring and evaluation.



Source: Williams (2018).

As discussed earlier, the Commission is recommending that the overarching principle of the Indigenous Evaluation Strategy be centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. This principle recognises that Australian government agencies need to draw on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people when evaluating policies and programs if outcomes are to improve. What this means in practice is sharing decision making, through effectively engaging with Aboriginal and Torres Strait Islander people on what to evaluate (evaluation priorities), evaluation questions, approaches, methods and practices, and the interpretation of evaluation findings.

5 Evaluating for quality results

Key points

- The quality of evaluation matters. Done well, evaluations provide robust and credible evidence that can be used with confidence to support better policy and program design. Done poorly, evaluations can produce misleading results.
- Building evaluation into the design of a policy or program is crucial for ensuring quality evaluation, and specifically that:
 - the data needed to evaluate are collected
 - Aboriginal and Torres Strait Islander people, and others affected by the policy or program are engaged early in the evaluation
 - the evaluation is fit-for-purpose and designed to maximise its usefulness.
- High-quality evaluations are also enabled by early planning and engagement, clear program logics and evaluation questions, rigorous and fit-for purpose evaluation methods, good data, ethical conduct, quality assurance measures and transparent evaluation methodologies, processes and findings.
- To achieve better outcomes, what Aboriginal and Torres Strait Islander people value, their expertise and lived experiences need to be reflected in how evaluation is undertaken and the outcomes policies and programs seek to achieve.
- There should be effective engagement with Aboriginal and Torres Strait Islander people on defining policy and program outcomes, evaluation questions, evaluation approaches, methods and conduct and on how evaluation findings are interpreted. This will improve the quality and use of evaluations and better align policies and programs with the needs and priorities of Aboriginal and Torres Strait Islander people.
- Mainstream policies and programs should routinely consider impacts on Aboriginal and Torres Strait Islander people — and evaluate those where the impact is considered significant. Early planning for evaluation is one way to ensure that the different effects of mainstream policies and programs on Aboriginal and Torres Strait Islander people are considered, and the data required to evaluate the effects are collected.

The quality of evaluation matters. Done well, evaluations can provide robust and credible evidence that can be used to improve policies and programs, and the lives of Aboriginal and Torres Strait Islander people. Done badly, evaluations can produce misleading results, leading to wasteful or potentially harmful policies and programs being continued. As the Organisation for Economic Cooperation and Development recently said:

... poor quality evaluations carry the risk of providing unfit evidence, or evidence that is subject to bias and undue influence. Poor quality evidence also implies that a policy that is ineffective, or even harmful, might either be implemented or continue to be. Finally, opportunities to use public funds more effectively may be missed. (OECD 2020a, p. 63)

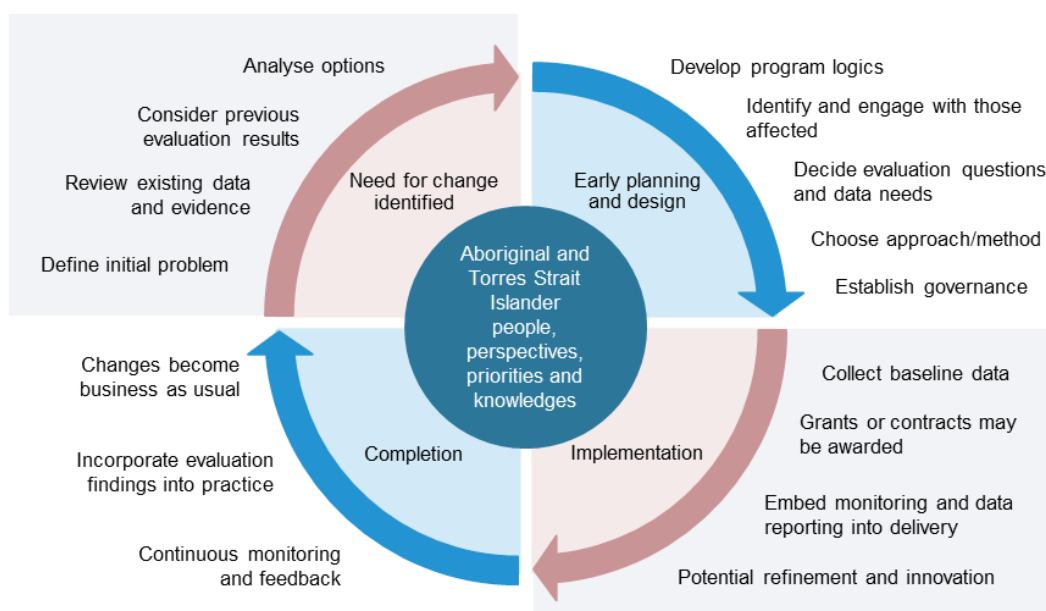
An important reason for an Indigenous Evaluation Strategy (the Strategy) is to ‘lift the bar’ on the quality of evaluations for policies and programs affecting Aboriginal and Torres Strait Islander people. This chapter looks at how to evaluate for quality results. Specifically, the chapter looks at:

- how to achieve quality in evaluation (section 5.1)
- engaging with Aboriginal and Torres Strait Islander people and their knowledges in policy and program design and evaluation practice (section 5.2)
- evaluating mainstream policies and programs affecting Aboriginal and Torres Strait Islander people (section 5.3).

5.1 What makes for a high-quality evaluation?

Evaluation is most effective when it is integrated into each stage of policy and program development, from setting policy objectives and collecting baseline data, through to using evaluation findings to inform future policy and program design (figure 5. 1).

Figure 5.1 **Evaluation should be considered at every stage of policy and program design and implementation**



Before a policy or program is fully formed, previous evaluation evidence should be reviewed and the evaluation evidence used to shape its design and the way it is implemented. As the policy or program is implemented, data and evaluation findings can provide feedback to strengthen the policy or program and/or address any unintended consequences. And

following the completion of an evaluation, the findings should support future policy and program development.

Building evaluations into the formative stage of policy or program design can ensure that policies and programs are designed so they can be evaluated (and evaluation designs and data collected are not limited). As Matthew James said at the Commission's Roundtable on *Better Indigenous Policies: The Role of Evaluation*:

The key to a good evaluation strategy is to build evaluation and analysis into policy design. ... if evaluation is not considered from the start many types of assessment will not be possible. Evaluation may not be possible at all unless it is built into the policy development budget. ... Another key reason that evaluation should be built into policy design is that it can be critical to have access to key data before a policy commences if that policy is to be evaluated well. (2013, p. 111)

The National Justice Project also said:

Evaluation must be considered as part of the overall policy cycle and not as a separate process. It is an essential component of policy and program design right from the start. This enables stakeholders, including Aboriginal and Torres Strait Islander peoples, to make informed decisions about the desired impact, outcomes and indicators of success. It also enables the incorporation of data collection into the planning, and the gathering of baseline data where necessary. (sub. 51, p. 2)

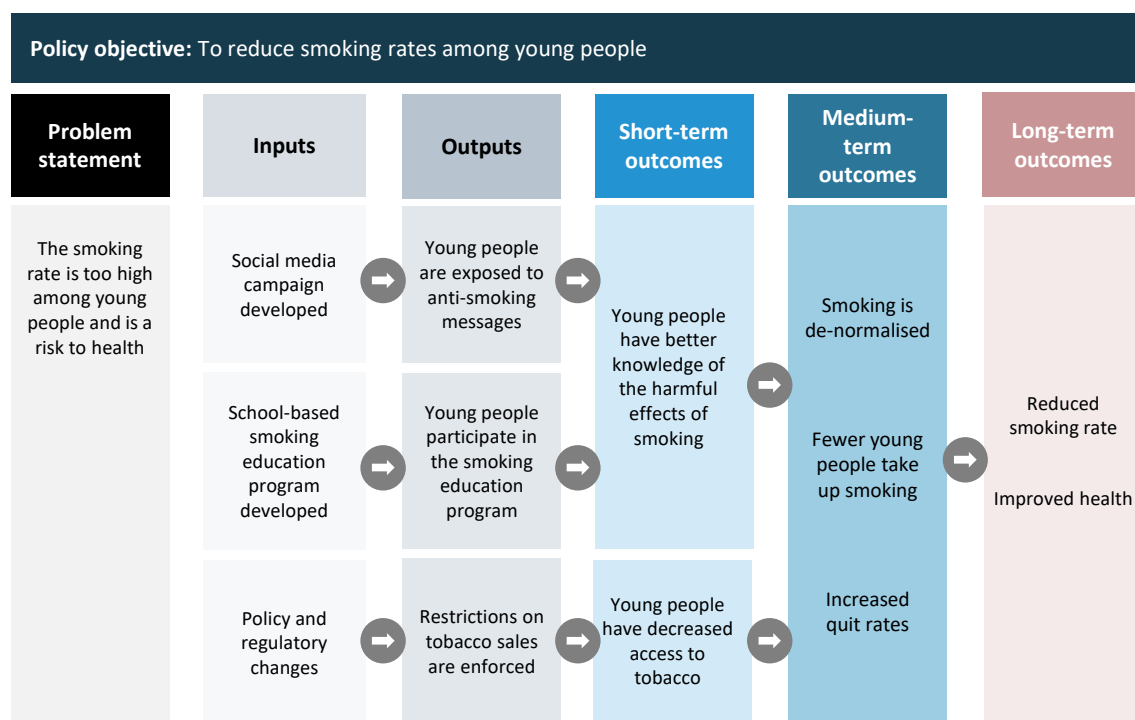
And on policy and program design, the Magenta Book notes that 'small changes in an intervention's design can make the difference between a high-quality, useful evaluation and one that is not able to answer the key questions under consideration' (HM Treasury 2020c, p. 13).

High-quality evaluation starts with an 'evaluable' policy

An important early step to a quality evaluation is understanding the policy or program and working out what to evaluate and what the evaluation questions should be. One way to be clear about the objectives of a policy or program is to develop a theory of change or logic model. A logic model describes the relationship between policy or program inputs, outputs and outcomes. It is a useful way to demonstrate the assumptions and evidence that the policy or program is based on and to identify the outcomes that need to be measured and the data required to measure the outcomes (figure 5.2 provides an example of a program logic for an anti-smoking initiative for young people).

Evaluation design centres on testing the program logic against actual outcomes delivered. Developing a logic model or theory of change can lead to both better policy planning and evaluations and can ensure that the things that really matter are measured during the evaluation (HM Treasury 2020c, pp. 24–26; OECD 2020a, p. 65).

Figure 5.2 Program logic of an anti-smoking policy for young people



Source: Adapted from Centre for Epidemiology and Evidence (2017, p. 11).

Early planning is key ...

Early planning is key to quality evaluations and is critical to ensuring the evaluation delivers useful findings (table 5.1 sets out questions that should be asked at the planning stage of an evaluation). It can also help build a culture of evaluation as evaluation becomes embedded in the policy cycle rather than being something that agencies do as an afterthought.

Building evaluation into the design of policies and programs helps ensure that the data and evidence needed to answer the evaluation questions and to produce credible results are collected (including, for example, the collection of baseline data and finding an appropriate comparison group). Early planning is also one way of ensuring that the potentially different effects of mainstream policies and programs on Aboriginal and Torres Strait Islander people, and other minority groups, are considered at the planning stage, and the data required to evaluate the effects are collected (section 5.3).

Table 5.1 Questions to consider at the planning stage of an evaluation

<i>The main steps</i>	<i>Questions to consider</i>
Define the policy objectives and outcomes to identify what to evaluate	What is the program logic or theory of change? Does it clearly state how the program outcomes are expected to be achieved and how the outcomes are linked to higher level outcomes? Do the intended outcomes align with the priority needs of Aboriginal and Torres Strait Islander people and other stakeholders?
Identify and engage with those who will be affected by the policy or programs	Who will be affected by the policy or program? (This should include those involved in delivering the initiative, users/recipients, others affected by the policy or program, and those interested in the evaluation results.)
Identify the evaluation questions	What do policy makers and those affected by the policy or program need to know? (Evaluation questions are best decided on with the main evaluation information users.)
Select an evaluation approach and method(s)	What approaches and methods will answer the evaluation questions? Is the evaluation a formative, process, impact evaluation, or some combination? Is an economic evaluation required? Are approaches/methods identified for each evaluation question? Are contextual factors that are likely to influence outcomes identified?
Identify data requirements (quality evaluations require good data)	What data are required? What is already being collected/available? If undertaking an impact evaluation, how will baselines be collected and when will impact be measured? What processes need to be put in place to allow the data to be collected? Has consideration been given to the design of data collection methods that are responsive to the needs, rights and cultural safety of respondents?
Identify ethics requirements	What ethics processes are required? Has time been allowed for ethics processes?
Identify resource and governance needs	What level of resourcing is required for the evaluation? Does the evaluation plan allow sufficient time and resources to engage meaningfully with relevant parties at each stage of the evaluation? What governance arrangements (a steering group or peer review) will need to be put in place? What quality control processes will there be?
Conduct the evaluation	Will the evaluation be conducted internally or be externally commissioned? Who will be responsible for tendering, project management and quality assurance? When will the evaluation start and end?
Use and disseminate the evaluation findings	What will the findings be used for? What decisions will they feed into? How will the findings be shared and disseminated? Will it be necessary to present the findings in different formats for different users? What approaches will be used to ensure information is considered during relevant policy decision making?

Source: Adapted from HM Treasury (2011, pp. 31–32).

The absence of good data, including baselines, was a commonly cited reason why evaluations were unable to answer evaluation questions (chapters 3 and 9). As the Western Australian Government said:

... it is important data requirements are documented in the evaluation plan and included in relevant contracts, and data sharing arrangements are finalised prior to the evaluation commencing. (sub. 74, p. 5)

And while over half of the agencies responding to the Commission's information request said that evaluation plans were 'always' or 'very often' required when developing new policies and programs, around a quarter said the evaluation plans were 'never' or 'rarely' developed when developing new policies and programs (appendix B).

... as is engaging with Aboriginal and Torres Strait Islander people so evaluations reflect their perspectives, priorities and knowledges

Developing an evaluation plan as part of the policy or program design enables Aboriginal and Torres Strait Islander people and others that are likely to be affected by the policy or program to be identified and involved early on in the process. Aboriginal and Torres Strait Islander people should play a role in shaping policy objectives, outcomes, evaluation questions, approaches and methods, data and data collection. As the World Bank said, 'the absence of stakeholder engagement in the definition of the 'evaluation problem' risks reducing the utility of an evaluation or even creating resistance to evaluation findings' (World Bank 2019, p. 12).

Planning early also means that thought can be given early on to how to reach and engage with particular communities or groups, how to engage in a culturally safe way, and ensure the time and resources are factored into the evaluation for such engagement.

Engaging with Aboriginal and Torres Strait Islander people and communities in policy design and at the various stages of evaluation ensures that what they value, their perspectives, and knowledges are taken into account (table 5.2). As many participants pointed out, this is critical if outcomes are to be improved (box 5.1). The National Mental Health Commission, for example, said that the Strategy should:

... incorporate engagement with people directly affected by a policy, program or service as a core principle of evaluation. This principle would encourage an approach to evaluation that moves beyond a simple focus on inputs and outputs, to a focus on experiences and outcomes. This approach would provide a deeper and more nuanced understanding of the impacts of policies, programs and services on the communities that are affected by them, and of the factors that drive positive outcomes. (sub 19, p. 2)

Table 5.2 Engaging at each stage of an evaluation

<i>Engagement with Aboriginal and Torres Strait Islander people is important during...</i>	
...definition of policy objectives and outcomes and what to evaluate	to ... • ensure the evaluation meets the needs of Aboriginal and Torres Strait Islander users • understand how the program logic or theory of change could look different for Aboriginal and Torres Strait Islander people compared to the rest of the population
... evaluation planning and as part of evaluation governance	to ... • determine inclusive evaluation objectives and questions • decide on appropriate evaluation design and methods • decide on monitoring and evaluation measures • plan for data collection • plan evaluation budgets and timeframes • plan engagement with Aboriginal and Torres Strait Islander evaluation participants • choose a suitable evaluation team • identify and respond to ethical issues
... data collection and analysis of evidence	to ... • embed effective data collection and monitoring arrangements (including with organisations who deliver services for government) • effectively track outcomes for Aboriginal and Torres Strait Islander people • enable timely feedback during implementation • incorporate relevant data, information and evidence • ensure evaluation involving Aboriginal and Torres Strait Islander people is conducted in a culturally safe way
... interpretation and reporting of results	to ... • accurately interpret and report data relating to Aboriginal and Torres Strait Islander people • validate or refine evaluation findings with evaluation participants in light of evidence • ensure that findings about Aboriginal and Torres Strait Islander people are presented accurately and respectfully
... use and dissemination of results	to ... • access, share and make use of evaluation findings • translate evaluation findings into policy and program improvements

Box 5.1 **Participants highlighted the importance of engagement**

Empowered Communities:

Include Indigenous people early in planning for and design of evaluation processes so that adaptive learning can be implemented during policy or program implementation and this information communicated to communities. (sub. 41, p. 11)

Catholic Care NT and the Australian Centre for Social Enterprise:

Understanding Indigenous life experiences and knowledges, which vary across communities and nations, should frame the development of outcomes and the consequent evaluation. (sub. 75, p. 6)

The Australian Evaluation Society:

Developing the theory of change, program theory or program logic with Indigenous people and communities helps to underpin the design and methods which are more culturally appropriate to the communities in which the evaluation is being conducted. ... Engaging with Indigenous people in co-design processes, would enable the factors that are important to them to be fully considered in the approaches and methods to be decided. (sub. 49, pp. 17–18)

Tony Dreise et al.:

Quality Indigenous evaluation requires program design where First Nations people have had a key role in determining clear objectives, articulating theories of change, and identifying relevant indicators of success. (sub. 33, p. 4)

Central Australian Aboriginal Congress:

Any evaluation of policies or programs affecting Aboriginal people must include the views and experiences of Aboriginal community members, leaders, and organisations and reflect the diversity of Aboriginal Australia with respect to gender, remoteness, culture, age, disability, sexuality etc. (sub. 48, p. 6)

Social Ventures Australia:

We recognise that best practice evaluation methods prioritise the involvement of the communities who are affected. (sub. 83, p. 4)

The Western Australian Government:

The critical factors that need to be considered are engaging and collaborating with Aboriginal and Torres Strait Islander people, and co-designing an appropriate evaluation framework. This will assist in building a long-term evaluation culture, and facilitate a sense of ownership by Aboriginal people in the process. (sub. 74, p. 4)

... technical rigour (including the elimination of bias)

High-quality evaluations require sound methodology, systematic assessment and rigorous inquiry. They also need to be conducted by evaluators who are technically and culturally capable (chapter 8). This means being able to apply systematic methods of inquiry that are appropriate to the evaluation in question to ensure credible results.

Evaluations need to use the most appropriate evaluation type, approaches and methods given the evaluation questions being asked, the data available and the time and resource constraints of the evaluation (chapter 4). The approaches and methods should also be culturally responsive and appropriate for the context. As the Organisation for Economic Cooperation

and Development said ‘choosing the appropriate evaluation method is paramount to an evaluation’s quality’ (OECD 2020a, p. 75).

Social Ventures Australia also emphasised the importance of rigorous methods:

We recognise that an emphasis on high quality, appropriate and rigorous evaluation methods and approaches are critical to generating evidence and data that accurately captures the experiences, aspirations and priorities of Aboriginal and Torres Strait Islander peoples. (sub. 83, p. 5)

Evaluations should also be replicable (this requires evaluation reports to systemically document all aspects of the evaluation, including evaluation methods, data, analyses, findings, limitations and conclusions).

Valid sampling designs and data analysis techniques are important for technical rigour. This includes ensuring appropriate sampling strategies and sample sizes (for a representative sample, to identify statistically significant change, and to achieve sufficient power in experimental designs, section 5.3). Establishing impact requires rigorous comparisons either between groups, in time or to alternative theories (HM Treasury 2020c, p. 16).

For qualitative evaluation, technical rigour relates to how evidence is gathered and analysed, and used to draw findings, and how information is reported (Anderson 2017; HM Treasury 2012; Rocco 2010). Qualitative studies are most rigorous when processes are systematic, clearly documented, transparent and auditable, when hypotheses are systematically tested, and when findings are reviewed and validated by participants.

Evaluation rigour can be enhanced by employing triangulation including using mixed method approaches (combining quantitative and qualitative methods, chapter 4).

It is important to be aware that evaluations can have biases that can compromise quality. As Patton said:

Our very process of taking in information distorts reality — all the evidence of social science indicates this. We have selective perception — some of us have rose-coloured glasses, some of us are gloom-and-doomers. We are not neutral; there is an emotional content to information. We need disciplined techniques to be able to stand back from that day-to-day world and really be able to see what is going on. We need approaches to help us stand back from our tendency to have biases, prejudices, and preconceptions. (cited in IRDC 2011)

Some of the forms of bias that can affect evaluation results include: confirmation bias (evaluators pay more attention to data or other information that support their pre-existing views); selection bias (underlying differences between a treatment and comparison group that result in them having different outcomes irrespective of the policy or program); and cultural bias (where evaluation findings are influenced by the evaluator’s own cultural values and perspectives).

Not addressing assumptions and biases when undertaking research involving Aboriginal and Torres Strait Islander people can result in a lack of trust by Aboriginal and Torres Strait Islander people in research findings. Regardless of the methods and approaches used,

evaluators and researchers need to be cognisant of their own biases when doing evaluation. As Laycock et al. said:

Western research approaches mostly aim to be unbiased when collecting and interpreting data. However, unbiased research is still built on the academic theory bases in which we work, our underlying hypotheses (what we think we will find), and the assumptions and values of the work discipline. Research observations, however direct, are filtered through the observer's worldview, which is influenced by race, culture, gender, how we have been raised and taught, and more. (2011, p. 10)

... objectivity

For evaluations to be credible, there needs to be a degree of objectivity. This can be achieved through independence of the process for conducting evaluations. For policy and program evaluation, several facets of independence are relevant:

- *Behavioural independence* — this is about the conduct, impartiality and integrity of the evaluator, the freedom the evaluator has to conduct the evaluation effectively (for example, having access to relevant data and information), and evaluator's capacity to work without coercive influence (related, for example, to their remuneration or employment).
- *Organisational independence* — this involves separating evaluation activities from policy and program management. It involves evaluation teams working separately from policy and program staff, either in separate units within the agency, or externally. It can also be functional, so that the managers responsible for making decisions on policies and programs are not also making decisions about evaluation funding, design and processes (OECD 2020a, p. 67; UN 2016, p. 11; WHO 2013, p. 2).

However, full independence in government evaluations presents some challenges. There may need to be close interaction between policy and program staff and evaluators at the evaluation design stage. The involvement of policy and program staff is also expected to enhance potential learnings about how a policy or program is working in practice. Also, commissioning evaluations to external parties does not fully guarantee independence, as commissioned evaluators can have an interest in complying with agencies' expectations to gain repeat business (chapter 8).

... good data

Robust and timely data are also critical for quality evaluation — without good data it is impossible to answer empirical evaluation questions (such as whether a policy or program has improved literacy skills or school attendance). As Lovett observed:

... unreliable and inappropriate data lead to the perpetuation of ineffective policies and programs because our ability to assess their outcomes and effectiveness is limited. (2016, p. 213)

The evaluation questions will determine what data need to be collected. Administrative and monitoring data can be important sources of data for evaluation, but new data often need to be collected for evaluations.

As discussed above, data requirements should be considered, and new data planned for, prior to any policy or program commencing. This will ensure baselines and counterfactual data can be collected, ethical and data protection issues are considered, and data costs are taken into account. Impact evaluations need to collect standardised data from both the treatment and control groups. And the sample size in any data collection needs to be large enough to deliver robust results. It is also important that evaluations correctly interpret the evidence collected.

One example of an evaluation found to be hampered by poor data is the Department of Social Services' evaluation of the Cashless Debit Card Trial (box 5.2).

Chapter 9 looks at data issues in more detail.

Box 5.2 Data collection concerns — an evaluation of the Cashless Debit Card Trial

The Cashless Debit Card Trial is an Australian Government initiative to test whether reducing the amount of cash available in a community reduces the overall harm caused by alcohol, gambling and drug misuse. It is issued to income support recipients and works like a regular bank card, except that it cannot be used to buy alcohol or gambling products, some gift cards, or to withdraw cash. Up to 80 per cent of a recipient's income support payment can be quarantined on the card.

In December 2014, the Australian Government announced it would roll out a trial of the Cashless Debit Card in Ceduna and the East Kimberley beginning in early 2016. Participation was mandatory for all working age income support recipients in the trial sites, but a large majority of participants were Aboriginal and Torres Strait Islander people. In February 2016, the Department of Social Services commissioned an external consultant, ORIMA Research, to evaluate the trial in the two sites. ORIMA completed the evaluation in August 2017, concluding that the trials were effective in reducing alcohol consumption and gambling, with some evidence of a related reduction in violence and harm.

In 2018, the Australian National Audit Office reviewed the Department of Social Services' implementation of the trial to determine whether it had been adequately monitored, evaluated and reported on. It found that arrangements were inadequate — there was a lack of robustness and rigour in data collection to measure impact, and the department failed to effectively use its own administrative data to assess any change in social harm. Also, the evaluation design did not examine cost efficiency or scalability, limiting its usefulness for informing further roll out of the trial. The Australian National Audit Office recommended that the Department of Social Services undertake a post-implementation review and cost-benefit analysis of the trial, improve its monitoring and evaluation arrangements, and continue to evaluate the Cashless Debit Card. The Australian Government is extending the operation of the Cashless Debit Card in existing sites to June 2021, and is transitioning to the Cashless Debit Card in other locations where income management is in place.

Sources: ANAO (2018); DSS (2020); ORIMA Research (2017).

... quality assurance and good governance

Quality assurance and assessment are also important for quality evaluations (HM Treasury 2020c, p. 75; OECD 2020a, p. 82). Quality assurance should occur throughout an evaluation so that evaluation design, planning and delivery are conducted systematically and conform to professional standards, and analytical standards are adhered to. Quality assurance can be conducted through peer review or steering committees. It can also take place after an evaluation is completed by an independent assessor, adding another layer of accountability.

The Department of Foreign Affairs and Trade (DFAT) has developed robust quality assurance standards and procedures to ensure quality monitoring and evaluation of its foreign aid programs (box 5.3).

Box 5.3 The Department of Foreign Affairs and Trade's evaluation quality assurance measures

The Department of Foreign Affairs and Trade (DFAT) uses several mechanisms to ensure quality monitoring and evaluation of its foreign aid programs.

- Program areas are responsible for initiating and managing a number of priority evaluations (agreed in an annual evaluation plan) each year, while DFAT's monitoring and evaluation standards define indicators of quality for each evaluation stage.
- DFAT's Office of Development Effectiveness (ODE) provides support to program areas to help them meet the monitoring and evaluation standards. The ODE also conducts a regular review of completed program evaluations to assess their quality.
- The ODE itself undertakes a number of strategic evaluations, which consider priority departmental policy, programming and development issues.
- DFAT's Independent Evaluation Committee oversees the progress of these evaluations, including reviewing the terms of reference and evaluation plan, considering regular progress updates, and reviewing interim and final reports. The Committee consists of three independent members.

Sources: DFAT (2017b, 2017c).

The OCHRE plan — Opportunity, Choice, Healing, Responsibility and Empowerment — is the NSW Government's policy to support strong communities in which Aboriginal people actively influence and participate fully in social, economic and cultural life. As well as Aboriginal governance arrangements, the evaluation of OCHRE included a key quality assurance role for Aboriginal communities involved in the evaluation. Evaluators presented draft findings and recommendations to communities that were involved in the program and the evaluation, with community feedback reflected in final reports. Evaluation reports were owned and controlled by participating communities, who were able to decide whether the reports should be publicly released and how the findings would be used (Ombudsman New South Wales 2019, p. 40).

... and transparency

Transparency of evaluations is also important for quality evaluation as it allows others to see whether an evaluation and its findings are valid, rigorous, and procedurally and methodologically sound. It also allows other users of evaluation (not just the commissioning agency or program area) to learn from the insights from evaluation and keeps governments and decision makers accountable for their responses to evaluations (chapter 7).

Transparency can also help to build trust and instils confidence with key stakeholders, which is important if their buy-in is needed to implement changes following evaluation (UN 2016, p. 12). Transparency in evaluation can also be considered an ethical requirement.

Transparent evaluation practice is maximised when:

- those affected by the policy or program (non-government organisations, communities and program users) are fully informed, engaged and empowered throughout the evaluation process, for example, by promoting awareness of the evaluation, actively seeking feedback on design, encouraging engagement, testing evaluation conclusions, and sharing findings
- evaluation reports contain detailed information explaining an evaluation's purpose, objectives, approaches, processes, sources of evidence, analysis, conclusions and limitations
- evaluation products (the report, but also the data and evidence supporting its findings) are released to relevant stakeholders or the public
- evaluation processes are subject to formal governance arrangements and external oversight
- evaluation processes are well documented and auditable (HM Treasury 2020c, p. 84; WHO 2013, p. 78).

External monitoring and review of agencies' evaluation practices can also provide incentives to improve quality (chapters 7 and 10).

Evaluations also need to be conducted ethically

Ethics in evaluation is concerned with good conduct — the conduct of evaluators, the integrity of the evaluation process, and the protection of participants. General principles for evaluation place obligations on evaluators to behave ethically in terms of:

- considering the need for an evaluation at the outset and seeking informed consent
- engaging appropriately and respectfully with participants in evaluation processes, including respecting rights, privacy, confidentiality, dignity and diversity of stakeholders and avoiding harm
- avoiding conflicts of interest in all aspects of their work

-
- systematic inquiry, and basing evaluation results on sound and complete information
 - reporting results in a fair and balanced way that acknowledges different perspectives (AEA 2018; AES 2013b; HM Treasury 2020c, pp. 76–79; Posavac and Carey 1997, pp. 85–101).

Ethical practice during evaluation should be guided by existing ethical guidelines for research involving Aboriginal and Torres Strait Islander people, including the Australian Institute of Aboriginal and Torres Strait Islander Studies' *Guidelines for Ethical Research in Australian Indigenous Studies* (2012) and the National Health and Medical Research Council (NHMRC) guidelines on *Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities* (NHMRC 2018b, p. 1). These frameworks emphasise the importance of ethical principles to good practice throughout the research process.

In the context of research undertaken with Aboriginal and Torres Strait Islander people and communities, the NHMRC said:

Ethical conduct of research is about making sure the research journey respects the shared values of Aboriginal and Torres Strait Islander Peoples and communities as well as their diversity, priorities, needs and aspirations; and that research is of benefit to Aboriginal and Torres Strait Islander Peoples and communities as well as researchers and other Australians. (NHMRC 2018a, p. 1)

For some evaluations, a formal ethics review will be required. Ethics review is 'a process of *prior* negotiation on points of ethics involving the intent and conduct of research and its results' (Fluehr-Lobban (1991, pp. 232–233), quoted in Davis and Holcombe (2010, p. 2)). Human Research Ethics Committees play an important role examining evaluation proposals to provide advice and recommendations on how to minimise ethical risks and promote ethical conduct in evaluation design and conduct.

While all evaluations should be undertaken ethically, not all require formal ethics review by an ethics committee (box 5.4). However, all evaluations should include a systematic assessment of ethical risks, and decisions on whether to request formal review by an ethics committee should be clearly documented.

Formal ethics review may not be required where evaluations:

- do not impose any risks to participants
- use existing data routinely collected by an organisation in the conduct of their work and for the purpose of quality improvement
- do not infringe the rights or reputation of the participants, providers or institutions
- use data that are not linked to individuals and does not violate the confidentiality of participant data (AIFS 2013; NHMRC 2014).

A number of participants pointed to the importance of ethical evaluation practice, while some also highlighted the time and costs of ethic review (box 5.5).

Box 5.4 When should ethics review for evaluation be sought?

The National Medical Health and Research Council (NHMRC) has guidance on ethical considerations specifically for evaluation and other quality assurance activities. This is in addition to other NHMRC guidance on broader research activities.

The NHMRC considers that it is up to institutions and organisations to establish ethics review procedures for evaluation. In doing this, the organisations' main objectives should be: (1) for evaluation participants to be afforded appropriate protections and respect; (2) for evaluation to generate outcomes to assess and improve service provision; (3) for those who undertake evaluation to adhere to relevant ethical principles and legislation; and (4) for organisations to provide guidance and oversight to ensure activities are conducted ethically, including a pathway to address concerns.

The NHMRC notes that there are many situations where, while some oversight of evaluation is required, ethics review is not necessary. This includes where:

- the data collected and analysed are coincidental to standard operating procedures with standard equipment and/or protocols
- the data are collected and analysed expressly for the purpose of maintaining standards or identifying areas for improvement in the environment from which the data were obtained
- the data collected and analysed are not linked to individuals.

The NHMRC recommends that ethics review be triggered in cases where:

- the activity potentially infringes the privacy or professional reputation of participants, providers or organisations
- data or analysis from evaluation activities will be used for another purpose
- information gathered about the participant is beyond that which is collected routinely
- there is testing of non-standard (innovative) protocols or equipment
- evaluation involves a comparison of cohorts, randomisation or the use of control groups or placebos
- evaluation involves targeted analysis of data involving minority or vulnerable groups whose data is to be separated out of that data collected or analysed as part of the main evaluation activity.

The NHMRC advises that organisations should develop guidance to determine when formal review from a Human Research Ethics Committee is required, and when lower, less formal review is appropriate, such as review by a low-risk committee, a sub-committee or individual member of a Human Research Ethics Committee, or another appropriate individual or body.

Source: NHMRC (2014).

Box 5.5 **Support for ethical evaluation practice, but some also highlight the costs**

Australian Evaluation Society:

The [Australian Evaluation Society] fully supports the ethical conduct of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. We propose ethical evaluation practice as a key principle of the Indigenous Evaluation Strategy. (sub. 49, p. 25)

Shelley Bielefeld:

It is not appropriate for evaluation projects that affect Indigenous peoples to be exempt from formal ethics review processes. Moreover these ethics review processes should be robust ... [It is recommended that] all policy and program design, monitoring and evaluation are conducted in line with Australian human ethics standards, particularly those specified in the Australian Institute of Aboriginal and Torres Strait Islander Studies' (AIATSIS) *Guidelines for Ethical Research in Australian Indigenous Studies*. (sub. 65, pp. 7–8)

Aboriginal Health Council of Western Australia:

Any research or data collection involving Aboriginal people [should] be subjected to rigorous ethics approvals processes. (sub. 42, p. 3)

Darling Downs and West Moreton Primary Health Network:

Unfortunately, many Ethical requirements are barriers to agile program development and implementation, slowing down processes and requiring many hours of work that take the project beyond the boundaries of intent. Often ethical committees met irregularly and if you miss a meeting or it is cancelled it can delay a project for up to six months. (sub. 6, p. 2)

The Department of Health (sub. 35, attachment C, p. 5) noted that the evaluation of the Aboriginal and Torres Strait Islander Australian's Health Program, which is national in scope, required at least eleven separate ethics applications, including multiple applications in three states, in order to meet stakeholders' expectations of a robust ethics review process.

Time and resources

Quality evaluations require time and resources. Agencies should determine early, as part of planning, the resources needed to undertake an evaluation that is robust, methodologically sound, and delivers on its objectives. Key considerations include the cost, time and data requirements that are needed to answer evaluation questions. As the Lowitja Institute said:

... there are wide-reaching benefits from successfully implemented evaluations. However, much of this benefit requires adequate resourcing to engage with stakeholders, to effectively plan and implement the evaluation, to disseminate the results, to support the uptake of learnings and new evidence into the system, and to monitor and evaluate the implementation of this and measure impact. (sub. 50, p. 9)

A number of participants argued that evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are often not adequately funded²⁵. The Central Australia Academic Health Science Network, for example, said:

... evaluation processes — even if in place — are often not resourced properly, or indeed for long enough, to produce rigorous results. (sub. 85, p. 3)

Meaningful engagement with Aboriginal and Torres Strait Islander people and more participatory forms of evaluation can require additional time and resources, compared to other evaluations. The Western Australian Government said this is often overlooked:

Engagement with Aboriginal and Torres Strait Islander people ... takes time and commitment to establish trusting relationships. This is often lacking in evaluation projects ... Poorly designed evaluations, which do not allow time or resources for meaningful engagement, can be interpreted as disrespectful and tokenistic by Aboriginal and Torres Strait Islander people, and generate a reluctance to participate. It should also be noted that English is not always a first language in some Indigenous communities. (sub. 74, p. 7)

The cost of an evaluation varies depending on the type of evaluation, its scope, its complexity, and the availability and accessibility of existing data. Specific factors that affect the costs of an evaluation include:

- *the engagement requirements of the evaluation* — including the breadth and depth of engagement needed, the geographic spread of the engagement, whether there are any cultural requirements or cultural protocols that need to be observed, and whether evaluation participants need support or compensation for their involvement
- *the data collection methods to be used* — including the extent that existing data will be used, whether new data need to be collected (and how they will be collected) and whether new indicators need to be developed
- *the level of ethics review required for the evaluation* — including the time and cost of obtaining formal ethics approval (if required)
- *the complexity of the evaluation* — including the number of questions the evaluation seeks to answer, and whether the evaluation will be multifaceted, or conducted in phases during implementation (for example, assessing both process and impact)
- *the amount of work already done that can feed into the evaluation* — including evaluation planning and preliminary data collection
- *the timeline of the evaluation* — more resources may be needed if the evaluation seeks to achieve a lot in a relatively short amount of time.

In response to the Commission's information request (chapter 3 and appendix B), Australian Government agencies supplied details on 207 evaluations. Information on the cost of evaluations was only available for two of these evaluations.

²⁵ For example, Children's Ground (sub. 73, p. 8), the Western Australian Government (sub. 74, p. 2) and the National Aboriginal Community Controlled Health Organisation (sub. 95, p. 5).

5.2 Engaging effectively with Aboriginal and Torres Strait Islander people and their knowledges

Quality evaluation relies on agencies engaging well with those affected by a policy or program. As discussed in earlier chapters of this paper, many of the concerns about the quality of evaluations were in the context of the lack of engagement with Aboriginal and Torres Strait Islander people — ‘somebody else tells our story for us’, and the lack of regard for Indigenous knowledges.

Engaging effectively with Aboriginal and Torres Strait Islander people and communities in policy and program design and evaluation is key to ensuring that Aboriginal and Torres Strait Islander people, their values, perspectives, knowledges, expertise and lived experiences are addressed and reflected in outcomes. The question is — what is effective or genuine engagement?

Hunt described effective engagement as:

... a sustained process that provides Indigenous people with the opportunity to actively participate in decision making from the earliest stage of defining the problem to be solved. Indigenous participation continues during the development of policies — and the programs and projects designed to implement them — and the evaluation of outcomes.

Engagement is undertaken with an understanding of the historical, cultural and social complexity of specific local or regional Indigenous contexts and with genuine attempt to share power in relationships that foster mutual trust. It requires adequate governance arrangements. (2013, p. 3)

Many participants spoke about engagement with Aboriginal and Torres Strait Islander people requiring relationships that are built on trust and respect and that allow Aboriginal and Torres Strait Islander people to engage as equal partners. Indigenous evaluation and research framework principles (table 5.3) also point to the importance of:

- relationships, reciprocity and shared responsibility
- Aboriginal and Torres Strait Islander people being engaged in each step of the evaluation process and as decision makers in evaluations
- working in ways that recognise Aboriginal and Torres Strait Islander people’s knowledges, cultures, languages, beliefs and identities.

Table 5.3 Principles from selected Indigenous evaluation and research frameworks

<i>Lowitja Institute Evaluation Framework for Aboriginal and Torres Strait Islander Health</i>	<i>NHMRC Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities</i>	<i>AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research (consultation draft)^a</i>	<i>Kaupapa Māori evaluation principles (New Zealand)^b</i>
• Partnerships • Shared responsibility • Engagement • Capacity building • Equity • Accountability • Evidence based • Holistic concept of health • Cultural competence • Data governance and intellectual property • Capitalising on Indigenous strengths	• Spirit and integrity • Cultural continuity • Equity • Reciprocity • Respect • Responsibility	• <i>Indigenous self-determination</i> • Recognition and respect • Informed consent • Cultural capability and learning • <i>Indigenous leadership</i> • Indigenous-led research • Indigenous perspectives and participation • Engagement and collaboration • Indigenous knowledge and data • <i>Impact and value</i> • Benefit and reciprocity • Impact and risk • <i>Sustainability and accountability</i> • Indigenous lands and waters • Ongoing indigenous governance • Reporting and compliance	• Importance of meeting with people face to face • Look and listen to develop understanding from which to speak • Collaboration and reciprocity • Be politically astute and culturally safe • Don't trample the mana (power, dignity and respect) of the people • Be humble in your approach

^a The 2012 AIATSIS Guidelines for Ethical Research in Australian Indigenous Studies list 18 principles under the broad categories of (1) respect, rights and recognition (2) negotiation, consultation, agreement and mutual understanding (3) participation, collaboration and partnership (4) benefits, outcomes and giving back (5) managing research: use, storage and access; and (6) reporting and compliance. AIATSIS will finalise a new Code of Ethics in 2020.

^b This table presents one interpretation of Kaupapa Māori evaluation principles, based on Cram (2009).

Sources: AIATSIS (2019); Cram (2009); Kelaher et al. (2018a); NHMRC (2018a).

Grey et al. discussed the concept of social learning and making ‘evaluation meaningful ... through principles that ‘indigenize’ the evaluation process including: respecting and valuing Indigenous knowledge and ways of knowing; building trusting relationships; putting people first; and strengths-based approaches to build capacity’ (2018, p. 86).

Effective engagement also requires evaluators to have the skills and experience to work in partnership with, and to draw on the knowledges of, Aboriginal and Torres Strait Islander people.

As discussed in chapter 1, in the context of what success of the Indigenous Evaluation Strategy would look like, a common theme in submissions was the importance of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in program and policy evaluation planning, design, implementation and priority setting. Many also called for the right to self-determination to be recognised in the Strategy, recognising that the principle of self-determination requires that Aboriginal and Torres Strait Islander people be involved in decisions that affect them, including the design, delivery and evaluation of policies and programs.

Findings from evaluations of successful programs also reinforce the importance of Aboriginal and Torres Strait Islander people and communities being engaged in the design, development and delivery of programs (chapter 3). The evidence also shows that what works is engagement through partnerships within a framework of self-determination and strategies that address power inequalities and genuine efforts to share power (Hunt 2013, p. 2).

Genuine partnerships with Aboriginal and Torres Strait Islander people can also strengthen evaluation capability and facilitate a sense of joint ownership. And it can strengthen trust and the cultural capability of non-Indigenous evaluators and those delivering policies or programs (box 5.6).

Measures of value should reflect Aboriginal and Torres Strait Islander people's preferences

Because evaluation is an exercise in assessing the value or merit of a policy or program (chapter 1), the metrics used need to take into account the measures of value that reflect the preferences of those directly affected by the policy or program.

Some participants suggested that evaluations often overlook the extent to which policies and programs represent value for Aboriginal and Torres Strait Islander people. The Office of Indigenous Education and Engagement, for example, said:

Mainstream programs are rarely assessed on their ability to deliver to the minority as well as the majority. If governments are looking to prioritise evaluation to achieve a better future for our First Nations peoples, then they need to think about the implications of the Indigenous population profile in combination with the policy evidence base. (sub. 98, p. 7)

And Patrick Sullivan et al., citing their work on value assessment in public services, said:

Our research shows that while government may assess the success of its programs by evaluating whether these programs have achieved their objectives, this is insufficient in terms of addressing whether public value has actually been created. To do this, the objectives must be aligned with a particular group's social and cultural values. (sub. 11, p. 1)

Other participants pointed out that considering cultural value in policy and program design is critical to improving Aboriginal and Torres Strait islander people's lives. Some cultural considerations include the extent to which a policy or program:

- maintains connections to Country, including Aboriginal and Torres Strait Islander people's spiritual, physical and cultural connections to land
- works in ways that are consistent with cultural and community values, norms and protocols
- respects Aboriginal and Torres Strait Islander people's rights to exercise cultural responsibilities, practices and expressions of cultural identity (including language)
- supports the continuity of the bonds and relationships that are important to Aboriginal and Torres Strait Islander people, including kinship, group and community connections
- demonstrates cultural capability and cultural safety when engaging with Aboriginal and Torres Strait Islander people (chapter 8).

Aboriginal and Torres Strait Islander evaluation frameworks are also helpful in understanding value from the perspectives of Aboriginal and Torres Strait Islander people (for example, Kelaher et al. (2018b, 2018a) and Williams (2001) as discussed in chapter 4).

An example of a program taking considering the importance of cultural value is provided in box 5.6.

The continuum of engagement

The literature refers to a continuum of engagement, which ranges from informing to participant-led²⁶ (figure 5.3). Engagement at the higher end of the continuum is considered important for more complex problems and policy areas where there is a need to work collaboratively to find policy solutions (Hunt 2013).

In this context, Dreise and Mazurski noted that:

... First Nations affairs policy is complex, due to a multitude of factors including history, colonisation, ideology, politics, race relations, geography, and socio-economic marginalisation. It is therefore reasonable to suggest that Western science and academic research are unlikely, by themselves, to provide a holistic picture or a complete understanding of this inherent complexity or of the pathways necessary to turn Aboriginal marginalisation around. (2018, p. 9)

²⁶ Also known as empowerment evaluation, as discussed in chapter 4.

Box 5.6 A qualitative evaluation of a Mobile Dialysis Truck

Rates of end-stage kidney disease among Aboriginal and Torres Strait Islander people in remote areas are disproportionately high. However, because haemodialysis is not offered in many remote areas, Aboriginal and Torres Strait Islander people have to leave their Country (with its traditions and supports) and relocate to metropolitan or regional centres to receive treatment. This disrupts the kinship and cultural ties that are important for their wellbeing.

The South Australian Mobile Dialysis Truck visits remote communities and allows patients living in regional or metropolitan centres to return home for significant community events, spend time on Country with family and friends and have dialysis on Country.

An evaluation of the Mobile Dialysis Truck involved face to face semi-structured interviews with 15 Aboriginal and Torres Strait Islander dialysis patients and 10 nurses who had attended trips across nine dialysis units. The interviews were conducted using a 'yarning' method (box 4.16) to collect information in a culturally safe way, and in a relaxed, informal setting. Realist evaluation methodology and thematic analysis were used to understand patient and nursing experiences with the Mobile Dialysis Truck.

Aboriginal and Torres Strait Islander patients said the consequences of leaving Country included grief and loss. The identified benefits of the service included: the ability to fulfil cultural commitments, minimisation of medical retrievals from patients missing dialysis to return to remote areas, improved trust and relationships between patients and staff, and improved patient quality of life. The Mobile Dialysis Truck was also found to be a valuable cultural learning opportunity for staff (with this form of health service potentially improving cultural competencies for nursing staff who provide regular care for Aboriginal and Torres Strait Islander patients).

The Mobile Dialysis Truck was found to improve the social and emotional wellbeing of Aboriginal and Torres Strait Islander people who had to relocate for dialysis. It was also found to build positive relationships and trust between metropolitan nurses and remote patients.

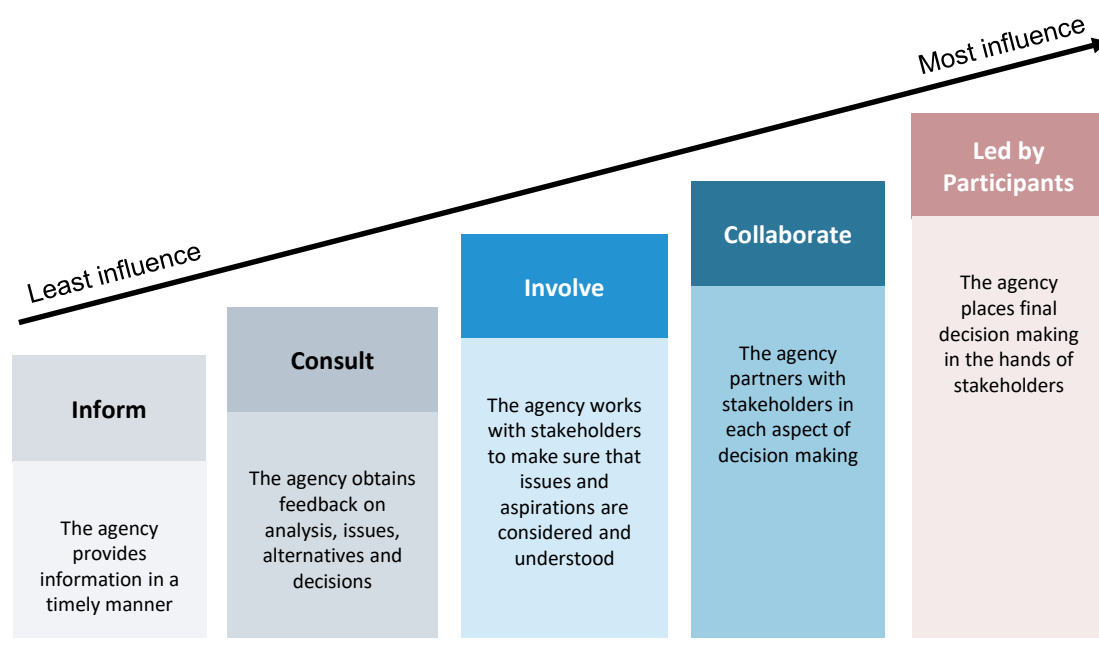
Some of the identified barriers to attending the Mobile Dialysis Truck on Country included infrequent trips, ineffective trip advertisement, poor patient health and a lack of appropriate or available accommodation for staff and patients.

The evaluators identified a number of potential areas for future research, including collecting:

- data from patients who had not been on trips, to elicit their experiences and compare data on quality of life
- quantitative data on hospital admissions and Royal Flying Doctor Service retrievals (to see if they were reduced in the presence of the dialysis truck)
- biomedical data on dialysis encounters (to see if adherence to dialysis and associated management regimes improved for patients who have spent time on the dialysis truck).

Source: Conway et al. (2018, pp. 1, 11).

Figure 5.3 A continuum of engagement



Source: International Association for Public Participation (2018).

Co-design

Many participants pointed to ‘co-design’ as an important approach to incorporating Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges into evaluations (box 5.7). While the term co-design is used and understood to reflect different levels of engagement across the continuum of engagement, it is usually more than consultation and falls within the collaborate and participant-led end of the continuum (figure 5.3).

It usually involves:

- a design-led process — iterative, person-centred and action oriented, and is about generating and testing new solutions at each stage of policy and program delivery
- participatory principles — enabling and empowering people with a direct interest in a policy or program to make decisions (this could be the deliverers of services with expertise in the field, or the users of services, whose ‘lived experience’ is a form of expertise)
- practical tools — using a range of methods to enable participants to talk about, enact or create solutions (Blomkamp 2018, pp. 732–733).

Box 5.7 **Participants' comments on co-design**

Tony Dreise et al.:

The design of evaluations should be a co-design process between government and First Nations peoples, ideally conducted as part of a program co-design process. Quality Indigenous evaluation requires program design where First Nations people have had a key role in determining clear objectives, articulating theories of change, and identifying relevant indicators of success. (sub. 33, p. 4)

Social Ventures Australia:

Delivering better outcomes for Aboriginal and Torres Strait Islander people requires a sophisticated and informed understanding on the appropriate measures and indicators of success. This is not only possible with Aboriginal and Torres Strait Islander peoples voice and participation in the design of, and decision making on, indicators of success. (sub 83, p. 3)

Central Australia Academic Health Science Network:

... Aboriginal and Torres Strait Islander peoples must have a role in designing evaluation processes at the macro and micro levels ... entrenching co-design, or preferably Aboriginal and Torres Strait Islander-initiation, of evaluation processes. ... For example, the Aboriginal Community Controlled Primary Health care sector would be charged with initiating and/or co-designing evaluations as to the effectiveness of government policies and programs in their sector, from funding models to governance issues ... (sub. 85, pp. 2–3)

Maggie Walter:

Real co-design in practice is an equal share in decision-making and equality of power. ... Aboriginal and Torres Strait Islander experts must be at the table as decision makers, and these representatives must have expertise in evaluation and data practice. Currently, some departments and agencies have Aboriginal and Torres Strait Islander people on advisory panels with good links throughout the Indigenous community, but who possess no expertise in evaluation. This somewhat negates their impact, meaning they can be shut down easily when it comes to technical issues. ... To legitimately participate in co-design, Aboriginal and Torres Strait Islander people need to have those skillsets in place. (sub. 112, p. 3)

In the context of the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, co-design is about engaging Aboriginal and Torres Strait Islander people and communities in the design and conduct of an evaluation in a way that is meaningful and engenders respect, empowerment and ownership (Dreise and Mazurski 2018, p. 8).

An example is the Department of Health's independent evaluation of the Australian Government's investment in Aboriginal and Torres Strait Islander Primary Health Care under the Indigenous Australians' Health Programme,²⁷ which is co-designed with experts

²⁷ The \$18.6 million evaluation is the largest investment in a single Indigenous health evaluation to date. It is being undertaken by a collaboration that includes Allen + Clarke, Monaghan Dreaming and the Centre for Indigenous Equity Research (Central Queensland University) (Australian Government Department of Health, sub. 35, Attachment A, pp. 1-2).

in Aboriginal and Torres Strait Islander health and with local Aboriginal and Torres Strait Islander communities. The evaluation's co-design mechanisms include:

- establishing a Health Sector Co-design Group to oversee the evaluation, bringing together health and evaluation experts (two co-chairs and more than half of the members are Aboriginal and/or Torres Strait Islander representatives)
- initial engagement during early co-design with six Aboriginal and Torres Strait Islander health organisations leveraging on existing relationships with the evaluation team
- further engagement with 103 members of various state and territory Aboriginal and Torres Strait Islander Health Partnership Forums, and with 36 other key organisations
- negotiating partnership agreements with primary health care providers and other stakeholders in 20 selected locations, who will work with the Department of Health throughout the evaluation process (Department of Health, sub. 35, Attachment A).

The Department of Health, reflecting on the evaluation, said:

Co-design provides a tool for system learning and adaptation and for Aboriginal and Torres Strait Islander Australians to have a real voice. It can be used to disrupt the status quo and help people to problem solve collectively and bring about positive change to improve outcomes for Aboriginal and Torres Strait Islander Australians. To date the evaluation team, the Department, and the Health Sector Co-design Group have reflected that co-design:

- Is a process
- Takes time
- Is about relationships
- Requires trust and trustworthiness. (sub. 35, Attachment A, pp. 7–8)

A co-design approach was also used to enable Aboriginal communities to engage in the New South Wales Government's OCHRE plan. The evaluation of OCHRE, which involved evaluators from the Social Policy Research Centre working in partnership with Aboriginal communities, enabled local community members to engage in evaluation design and to articulate their own measures of success (Aboriginal Affairs NSW 2018). The co-design process involved:

- informing communities of the opportunity to co-design the evaluation — the evaluation team spoke with local people and regional staff to raise awareness, support and buy-in of evaluation co-design
- conducting co-design workshops — evaluators held a series of workshops in communities to talk about the evaluation and seek advice on how it should be conducted
- agreeing the outcomes of co-design — obtaining formal agreement of the community on the decisions made in the co-design workshops including the overall design of the evaluation, and what the local community viewed as success and how it might be assessed (Dreise and Mazurski 2018, pp. 21–25).

Published evaluation plans, that have been agreed with each participating community, outline who in the community will be involved in the evaluation, how data will be collected, how local Aboriginal community researchers will be engaged, data collection methods, and how data will be analysed and reported.

Another example of co-design with Aboriginal and Torres Strait Islander people is described in box 5.8.

Box 5.8 A co-designed evaluation of a school-based trauma healing program

The evaluation of the intergenerational trauma project delivered by the Aboriginal and Islander Independent Community School (Murri School) used a co-design process. The four-year partnership between the Murri School and the Healing Foundation involved engaging with school students, their families and school staff to inform the future direction of the project. It included:

- a series of yarning circles^a, facilitated with students, families and school staff. The questions to guide the yarning circles were developed in partnership with Aboriginal and Torres Strait Islander staff from the school to ensure they were appropriate, and staff led the yarning circles to ensure a culturally safe environment. The circles sought insights into what factors cause stress for kids and families and what makes families strong
- a reflective practice circle with the Murri School Healing Team which involved identifying program strengths as well as priority areas for further focus and professional development
- a program logic workshop to allow the team to revise project goals, outcomes activities and outputs (based on their own experiences and input from students, families and the wider school staff). The Healing Foundation commented that ‘while program logic is a western planning concept, it is a useful tool when adapted to allow Aboriginal and Torres Strait Islander ownership of program design, goals and measures of success’.

^a ‘Yarning’ is a Aboriginal and Torres Strait Islander research method and cultural practice that uses story telling as way of learning and for knowledge exchange (chapter 4).

Source: Healing Foundation (2017).

Some of the potential benefits of co-design, as identified by the Department of Health, include that it:

- enables Aboriginal and Torres Strait Islander groups to help shape thinking and inform decision-making processes
- supports ongoing improvements to cultural safety and cultural appropriateness
- enables better relationships and cooperation between key stakeholders
- increases levels of support and enthusiasm for innovation and change
- improves knowledge of user needs and understanding of how services impact them
- enables testing of current understanding and emerging ideas in real time with relevant parties (DoH 2018).

Co-design can also improve transparency — participants are aware of (and engaged in deciding on) evaluation questions, evaluation approaches and methods and how to disseminate evaluation results. And because co-design is collaborative in nature, as Bradwell and Marr (2008, p. 17) said, ‘[it] shifts power to the process ...’. Co-design also recognises that evaluation is closely connected to the design, purpose and implementation of a policy or program itself.

However, co-design can be resource and time intensive. A reflection from the NSW Government OCHRE evaluation was that the co-design process was complex for evaluators and participants alike, and that it was hard to gauge ahead of time the resources and time required for co-design (Dreise and Mazurski 2018, pp. 24–25).

Some of the other insights from the OCHRE evaluation were that:

- community members need to understand what co-design is, and what may be expected of them as they take part in it (asking communities to participate in co-design can be asking for a significant commitment)
- community members need to be equal partners and be given relevant information about the evaluation theory and practice. They also need to have a good understanding of the initiative being evaluated, and have opportunities to co-design the evaluation
- evaluation teams need to become familiar with cultural protocols and follow them (this is important for establishing the community’s trust in the evaluation and the staff involved)
- flexibility is critical — evaluation teams need to be able to work with a community according to the community’s need (Dreise and Mazurski 2018, p. 17).

In summary, governments across Australia are embracing the idea of co-design in their work with Aboriginal and Torres Strait Islander people. For example, the Joint Council on Closing the Gap — which is made up of Australian, state and territory governments and representatives of Aboriginal and Torres Strait Islander peak organisations — agreed that one of their reform priorities would be:

Developing and strengthening structures to ensure the full involvement of Aboriginal and Torres Strait Islander peoples in shared decision making at the national, state and local or regional level and embedding their ownership, responsibility and expertise to close the gap. (JCOCTG 2019b, p. 1)

However more research and experience is needed to understand how co-design works in practice, both in policy and program design and in evaluation itself. As Blomkamp (2018, p. 734) said, ‘while co-design may have transformative effects, many of the claims about its benefits have not been rigorously evaluated’.

Governance arrangements for engagement

Aboriginal and Torres Strait Islander representation on evaluation governance or steering committees is another way of ensuring that the perspectives and experiences of Aboriginal and Torres Strait Islander people are incorporated into evaluation design and processes.

Evaluation governance or steering committees are usually responsible for overseeing the whole evaluation process, ensuring the quality of an evaluation, and that it delivers on its intended objectives. The main responsibilities of evaluation governance committees include:

- ensuring that the needs of decision makers and those affected by the policy or programs are addressed in the evaluation
- seeking agreement on the nature and scope of the evaluation among senior departmental staff, program managers and the evaluation team
- approving the evaluation strategy and endorsing the terms of reference
- making decisions on evaluation conduct and design
- ensuring the views of all those affected by the policy or program and evaluation users are considered
- monitoring the progress of the evaluation to ensure that it is on track, accords with the evaluation terms of reference and, where appropriate, approving modifications to the terms of reference during the course of the evaluation
- reviewing the overall quality of the evaluation presenting the findings of the evaluation and its recommendations to relevant parties for follow-up action (Department of Finance 1994a, pp. 37–38).

The National Indigenous Australians Agency's Indigenous Evaluation Committee provides Aboriginal and Torres Strait Islander oversight of evaluation activities within the Agency. The Committee is made up of three Aboriginal and Torres Strait Islander members and a non-Indigenous member with extensive evaluation experience. It is responsible for ensuring the conduct and prioritisation of evaluations of programs and activities under the Indigenous Advancement Strategy are impartial and support transparency and openness in the implementation of the Agency's Evaluation Framework. The Committee also assists in bringing the diverse perspectives of Aboriginal and Torres Strait Islander people into the evaluation process (DPMC 2018c).

Local engagement strategies

Relationships that have a local focus and enable local decision making can be important for place-based initiatives, community-led projects, and for policies and programs delivered in remote Aboriginal and Torres Strait Islander communities. The Generation One, Minderoo Foundation, for example, said that locally developed evaluation strategies should be 'driven by key community representatives who know the community needs and aspiration' (sub. 5, p. 5).

Referring to service delivery in remote and discrete Aboriginal and Torres Strait Islander communities in Queensland, the Queensland Productivity Commission also noted that an effective evaluation framework should:

... enable and foster local decision-making, and encourage adaptive practice (including learning from failure). For this to occur, evaluation, including reporting and compliance activities, must support the needs of communities rather than just government. This is not to say that an evaluation framework does not need to provide accountability for taxpayer's money — it does — but there should be greater consideration of how evaluation can improve information provision to the communities, individuals and stakeholders that are best placed to make day-to-day decisions about how services are delivered. (2017, p. 236)

Often, locally-led evaluation requires relationships to be built over the long term — not just for the life of a single evaluation — and throughout policy and program implementation. This builds trust within communities, maximises engagement, and enables stakeholders to build evaluation capacity over time. Some examples of locally-based engagement for evaluation are provided in box 5.9.

The Commission is interested in hearing participants' views on what effective engagement models for policy and program evaluation look like, including information on when co-design is most useful and examples of co-design in practice.

INFORMATION REQUEST 5.1

The Commission is seeking information on effective engagement strategies for evaluation. What engagement models are most effective? For what types of evaluations is co-design most useful? Why?

5.3 Evaluating mainstream policies and programs

As discussed in chapter 1, the majority of Australian government spending for policies and programs affecting Aboriginal and Torres Strait Islander people is on mainstream initiatives. For mainstream policies and programs, Aboriginal and Torres Strait Islander people can:

- represent a small segment of the overall population (examples include Medicare and early childhood education)
- be over-represented among the user population because of factors such as targeting, geographic location, or relative need (examples include housing assistance, the Community Development Program, and youth justice programs).

Box 5.9 **Some examples of local engagement**

Yuwaya Ngarra-li

Yuwaya Ngarra-li, which means ‘vision’ in the Yuwaalaraay/Gamilaraay language, is a partnership between the Dharriwaa Elders Group (DEG), Walgett, and the University of New South Wales. Its purpose is to improve the wellbeing, social, built and physical environment, and life pathways of Aboriginal people in Walgett through collaborating on evidence-based programs, research projects and capacity building. The DEG has a membership board with clear rules for how it works with its members, community and affiliates. There is a high level of trust and confidence in the DEG among the Walgett community and in the work of *Yuwaya Ngarra-li* by association.

GroundUP

GroundUP undertakes both research and service delivery simultaneously in Aboriginal townships of the Northern Territory. This enables collaborations in which services are continuously designed, delivered and evaluated. Place-based and site-specific monitoring and evaluation is:

- carried out under the guidance of Aboriginal Elders
- developed through processes of collaborative design, with Elders, co-researchers and other authorities, as well as government, non-government and Aboriginal organisations
- oriented around measures and principles of success emergent in processes of co-design
- connected to the development of local Indigenous researcher capability, and opportunities for recognition and employment of Indigenous researchers.

Ground Up recently completed a project with the Northern Territory Government and local Aboriginal Elders, researchers and community members to design processes for evaluating government initiatives in several remote communities. The first step was developing small local research teams (comprised of experienced researchers and mentorees) who could design and carry out community consultations. The second step was for these teams to conduct real-time evaluations of government engagement, providing feedback to government on their engagement practices. As a result, proposed evaluation engagement protocols have been developed with several communities.

Empowered Communities

Empowered Communities (EC) was launched in 2015 as a reform initiative and a new way for the Australian Government to work with Aboriginal and Torres Strait Islander communities. EC has a place-based focus — eight regions are represented by eight Aboriginal and Torres Strait Islander-led ‘backbone’ organisations that coordinate local initiatives in each region. EC has been designed as a long term initiative, with regions embedding a ‘learn and adapt as you go’ approach that sees developmental evaluation undertaken from the start of program implementation. This supports innovation and quickly applying lessons learnt as implementation progresses. It also allows Aboriginal and Torres Strait Islander leaders and the Government to share responsibility for identifying what is working, what is not and why, and adjusting actions accordingly. The EC backbone organisations are supported by an overarching monitoring, evaluation and adaptation framework that has been co-designed with the National Indigenous Australians Agency, but each region tailors this framework to its own individual circumstances and priorities.

Sources: Charles Darwin University and NT Government (2017); Empowered Communities (2015); GroundUP, Northern Institute and Charles Darwin University (sub. 62); NIAA (nd); *Yuwaya Ngarra-li* partnership, Dharriwaa Elders Group and UNSW (sub. 38, attachment).

A number of participants raised concerns about evaluations of mainstream policies and programs not considering the impact on Aboriginal and Torres Strait Islander people. The Western Australian Government, for example, said:

Too often, the impact of mainstream policies and programs on Aboriginal and Torres Strait Islander people is not adequately considered in evaluation. This is particularly of concern, as 'four in every five dollars spent by the Australian Government in providing services to Aboriginal and Torres Strait Islander people is spent through mainstream programs and services'. (sub. 74, p. 2)

Responses to the Commission's information request also showed that about two thirds of the evaluations of mainstream policies and programs did not consider the effects of policies and programs on Aboriginal and Torres Strait Islander people. And some of the mainstream policy and program evaluation reports provided to the Commission also noted that an analysis of impacts for Aboriginal and Torres Strait Islander people was not undertaken (or was limited) because of a lack of data (box 5.10 provides an example of an evaluation that was not able to answer questions about impact because of a lack of data).

Aboriginal and Torres Strait Islander people can have different experiences to other Australians with mainstream policies and programs. For example, Aboriginal and Torres Strait Islander people living in remote areas may be less able to access services than the rest of the population. They can also face barriers to accessing mainstream policies and programs (box 5.11), for example, the programs may not be culturally safe or affordable, or they may not be aware of programs available to them (AHRC 2012; Li 2017; Victorian Auditor-General 2014). And even when Aboriginal and Torres Strait Islander people access programs at the same rate as other Australians, mainstream initiatives can mean poorer outcomes for Aboriginal and Torres Strait Islander people if service practices do not address differences in need.

Box 5.10 Does HIPPY benefit Aboriginal and Torres Strait Islander children?

The Home Interaction Program for Parents and Youngsters (HIPPY) is a two-year home-based early childhood enrichment program aimed at supporting parents in their role as their child's first teacher. Brotherhood of St Laurence (BSL) has an exclusive licence to deliver HIPPY and operates in 100 sites nationally with funding from the Department of Social Services.

Although a mainstream program, the Australian Government invested in HIPPY to improve early learning outcomes for Aboriginal and Torres Strait Islander children. It committed more than \$100 million from 2008 to 2017 to continue HIPPY, and to expand the program to 50 more locations with a focus on Aboriginal and Torres Strait Islander communities (DSS 2015).

In 2010, BSL was commissioned to conduct an interim evaluation of HIPPY in partnership with Monash University (Liddell et al. 2010, 2011). The evaluation report provided some indicators showing an improvement to children's school readiness after one year, on average, relative to baseline data and a comparison group; however, the data excluded equivalent measures for Aboriginal and Torres Strait Islander children. Instead, the evaluation drew from qualitative reports on the progress of implementation at five Aboriginal and Torres Strait Islander-focused HIPPY sites. The reports noted that engagement with families was a challenge, but parents' informal feedback was positive, suggesting that HIPPY held 'significant promise' for Aboriginal and Torres Strait Islander families (Liddell et al. 2011, p. 108).

In 2018, ACIL Allen Consulting was commissioned to evaluate HIPPY (ACIL Allen Consulting 2018). The methods used included a literature review on the efficacy of the HIPPY model, an examination of program data, and consultations with 20 HIPPY providers. Based on 2014 and 2015 cohort data, the experiences of Aboriginal and Torres Strait Islander children and families were generally positive and similar to the overall population of families. Lower program uptake and completion was evident in remote Aboriginal and Torres Strait Islander communities, but the evaluation did not examine why this occurred. The literature review found limited evidence on the provision of HIPPY to Aboriginal and Torres Strait Islander families, and the evaluation report recommended that, as a high priority, the Department of Social Services collect data and continue to evaluate the program for Aboriginal and Torres Strait Islander children.

Neither evaluation answered the question of whether the HIPPY program improved learning readiness of Aboriginal and Torres Strait Islander children. While departmental program documents and the BSL evaluation report indicate that an Indigenous-specific evaluation was initially planned, no evaluation has been released to date.

Sources: ACIL Allen Consulting (2018); DSS (2015); Liddell et al. (2010, 2011).

Box 5.11 **Barriers to accessing mainstream government services**

The concept of ‘access and equity’ in service delivery relates to the extent to which government policies and programs can meet the diverse needs of people and achieve equitable outcomes. Potential barriers to access for Aboriginal and Torres Strait Islander people include:

- *lack of cultural safety* — policies and programs are delivered in ways that are insensitive or disrespectful toward Aboriginal and Torres Strait Islander people’s cultural identities, needs and practices
- *lack of awareness* — people do not know what services are available in the first place
- *language and literacy* — access to information or service delivery can be inadequate for some Aboriginal and Torres Strait Islander people who speak English as a second language or who have lower English literacy skills. A lack of access to Aboriginal and Torres Strait Islander language translators can be a barrier
- *lack of access to local services or transport to services* — people simply cannot get to the service location. This may be particularly relevant for Aboriginal and Torres Strait Islander people living in regional or remote areas. Access problems may also relate to digital or phone services
- *racism and discrimination* — when Aboriginal and Torres Strait Islander people are not treated with the same respect, or receive a lower quality of service, compared to other Australians
- *shame, embarrassment and fear* — the policy or program is delivered in ways that stigmatise people or causes distress
- *administrative complexity* — administrative processes are burdensome and difficult for people to navigate
- *mobility* — people who move frequently or who are without stable accommodation may not have proof of identification or other personal records on hand to qualify for assistance. The continuity of services may also be compromised if people are mobile
- *lack of affordability* — the costs of services are too high for people on low incomes to afford
- *inadequate servicing* — policies and programs do not cater for people with higher or different kinds of need. For example, services do not provide the outreach or additional support needed to ensure all eligible people receive the service.

Source: Adapted from the Victorian Auditor-General (2014).

Early planning for evaluation is one way to ensure that 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. Important evaluation questions for mainstream policies and programs are:

- Is it working differently for Aboriginal and Torres Strait Islander people, or differently for groups of Aboriginal and Torres Strait Islander people (or other groups)? Why, or why not, might it be working differently for different groups?
- Are Aboriginal and Torres Strait Islander people accessing the policy or program at the same rate as other Australians? If not, why not?
- Are there barriers to Aboriginal and Torres Strait Islander people accessing the policy or program? Is the policy or program being delivered in a culturally appropriate way?

The Western Australian Government suggested that the Strategy should require mainstream policies and programs to consider, and include, a comparison of achieved outcomes for Indigenous and non-Indigenous people in evaluating success on the basis that this will:

- capture essential information about changes in the social and economic wellbeing of Aboriginal and Torres Strait Islander people
- highlight to mainstream providers the need, as part of their core business, for programs to consider the impact of program efficacy for Aboriginal and Torres Strait Islander people
- create an important evaluation in itself — to measure the changing appreciation by Australian Government agencies in considering the impact and success of their programs on Aboriginal and Torres Strait Islander people. (sub. 74, p. 2)

As discussed in chapter 7, the Commission is proposing that, under the Indigenous Evaluation Strategy, all New Policy Proposals that are expected to have significant and/or differential effects on Aboriginal and Torres Strait Islander people are to include an Indigenous Evaluation Threshold Assessment. This will ensure that these new policies and programs will have a preliminary evaluation plan and budget in place.

The approach to evaluating the effects of mainstream policies and programs on Aboriginal and Torres Strait Islander people should 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, as well as what is already known about the impact of the policy or program on Aboriginal and Torres Strait Islander people. Focus groups or case studies could examine particular issues affecting Aboriginal and Torres Strait Islander people more closely. A targeted evaluation could also be undertaken to look specifically at the impact of a mainstream policy or program on Aboriginal and Torres Strait Islander people.

Small sample sizes

An issue for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people is small sample sizes. Many data sources cannot be used for evaluation because they do not have sufficient numbers of Aboriginal and Torres Strait Islander people. Small sample sizes can also limit statistical power. As Deborah Cobb-Clark said:

... many data sources are unsuitable for Indigenous program evaluation because they do not have sufficient numbers of Indigenous respondents for analysis. Even when quantitative analysis is possible, small sample sizes can drastically limit statistical power. This means that, given realistic sample sizes, only very large program impacts are likely to be detected at standard statistical levels. (2013, p. 85)

One way to address small samples is to ‘oversample’ — that is — increasing the number of Aboriginal and Torres Strait Islander people in sample data so that it exceeds proportionate representation within the reference population (box 5.12). Oversampling improves the accuracy of results by increasing statistical power, and reduces statistical bias that can be problematic when working with small numbers.

Box 5.12 An example of oversampling — the Adult Admitted Patient Survey

The Adult Admitted Patient Survey asks for feedback from people who have recently been admitted to a New South Wales public hospital. Each year the survey is sent to between 70 000-85 000 people approximately three months after the end of their hospital stay. Participants are randomly selected to complete the survey in order to provide a representative sample across New South Wales.

For the 2014 survey, Aboriginal and Torres Strait Islander patients were oversampled (sampled at a rate higher than the proportionate representation in the total patient population) to ensure that a sufficient number of responses were collected to allow statistical analysis.

The oversampling strategy involved the selection of all eligible Aboriginal and Torres Strait Islander patients (i.e. census design). Although Aboriginal and Torres Strait Islander people made up 2 per cent of the population, they represented 10 per cent of survey respondents. (The response rate among Aboriginal and Torres Strait Islander patients was 21 per cent, which was much lower than the response rate for all patients (44 per cent), and was another reason to oversample.)

Responses were then weighted to ensure that the overall survey response rate reflected a response rate that would be observed if patients were sampled proportional to the actual patient population.

Sources: Bureau of Health Information (2015, 2016).

Oversampling can also counteract problems with low response rates and participant retention that can make it difficult to maintain sample size over the course of a policy or program (especially if some small participant groups drop out at higher rates).

Not oversampling is one reason for why evaluation datasets are often of limited use for examining the effects of a policy or program on smaller populations, such as Aboriginal and Torres Strait Islander people. However, oversampling can be expensive and it will not always be possible, especially in cases where a policy or program does not have a critical number of Aboriginal and Torres Strait Islander participants in the first place. Data collection and sampling strategies are discussed in detail in chapter 9.

It may also be possible, if a policy or program is run in several places, to aggregate results to increase the sample size (that said, care needs to be taken to ensure implementation and location are similar enough to ensure results are comparable).

Most evaluations will require some sort of sampling strategy across individuals, minority groups, sites or time periods. What is important is asking questions early about what data and sampling strategies are required to measure impacts for Aboriginal and Torres Strait Islander people. Considering sampling issues early means appropriate sampling strategies are chosen and data are collected.

Another way to address small sample sizes is to introduce secondary methods (such as case studies or other qualitative methods) to provide more detailed information for smaller

populations. One example is a recent evaluation of the National Disability Insurance Scheme (Mavromaras et al. 2018). Recognising that people with a disability living in regional and remote areas can face additional challenges to those people living in metropolitan areas, the evaluation included a case study in the Northern Territory, with a focus on the Barkly trial site (box 5.13).

Data collection, analysis and reporting should also recognise diversity and heterogeneity among Aboriginal and Torres Strait Islander people and be cautious about making broad generalisations. For national or state-wide initiatives, important differences may be observed when data on Aboriginal and Torres Strait Islander populations are analysed by urban, regional and remote areas, and by gender, age, education, and socio-economic status.

Box 5.13 National Disability Insurance Scheme in a remote location

An evaluation of the National Disability Insurance Scheme (NDIS) included a case study on the operation of the NDIS in a remote location, the Barkly region in the Northern Territory.

The study of the NDIS trial in the Barkly region focussed on seven themes: the supply and demand of disability support services; the disability sector and its workforce; choice and control (including self-management); reasonable and necessary supports; participation, wellbeing and aspirations; fairness, equity and access; and the interface between the NDIS and mainstream sectors.

In-depth qualitative interviews were conducted across two waves with NDIS participants and their families and carers, non-NDIS participants, disability service sector representatives, and National Disability Insurance Agency managers and staff.

Some of the findings were that:

- prevalence of disability and associated health and social support needs was high in the area. However, this did not seem to match the number of actual participants in the NDIS
- understanding of the NDIS and its processes was low throughout the course of the evaluation, with awareness greater among non-Indigenous than Aboriginal participants
- outcomes for NDIS participants were variable, with better outcomes for people with good English literacy and/or computer skills and those with strong advocates
- NDIS participants living in remote Aboriginal communities had particularly limited outcomes under the NDIS with a lack of adequate funding for transport, respite, support coordination and skill development reported
- choice and control remained an almost theoretical concept with these processes challenged by the virtual absence of a viable disability support market
- some improvements in social and economic participation were reported, but further progress was stymied by limited local services and opportunities.
- while attempts were made over time to adapt NDIS processes to meet local need, the approaches were considered ineffective. The NDIS was not perceived to have adapted sufficiently to address the specific needs of the Barkly region. A model which was more culturally sensitive and appropriate to remote needs and service delivery was recommended.

Source: Mavromaras et al. (2018).

6 What to evaluate

Key points

- It is not practical or feasible to rigorously evaluate all policies and programs. Setting priorities for evaluation is critical to ensuring that the limited resources Australian Government agencies have are directed to where evaluation can best inform the development of policies and programs to improve the lives of Aboriginal and Torres Strait Islander people.
- Currently an integrated whole-of-government approach to prioritising evaluations across Australian Government agencies does not exist. This means there are different approaches across agencies for setting evaluation priorities. There is also a lack of transparency and accountability around how decisions are made about what is evaluated. And engagement with Aboriginal and Torres Strait Islander people on identifying and determining priority areas for evaluation is limited.
- A set of priority evaluation areas aimed at improving the lives of Aboriginal and Torres Strait Islander people should be formalised to provide guidance to agencies on what to evaluate under the Indigenous Evaluation Strategy — that is, beyond agency-specific considerations. These priority areas should be decided in partnership with Aboriginal and Torres Strait Islander people.
- A criteria-based priority setting process should be adopted by each Australian Government agency to determine which of its new and existing policies and programs (in scope of the Indigenous Evaluation Strategy) should be prioritised for evaluation, and to determine the extent to which they are evaluated. Priorities should be determined based on policy and program impact, risk profile, strategic significance and expenditure, as well as on Aboriginal and Torres Strait Islander people's priorities.
- The process of prioritising and undertaking evaluations should be transparent. All departments, and major agencies (as deemed appropriate by departments), should centrally coordinate and publish, on an annual basis, a rolling Three Year Evaluation Forward Work Plan. The plans should detail:
 - the scope of key policies and programs that contribute to government-wide evaluation priorities aimed at improving the lives of Aboriginal and Torres Strait Islander people
 - the methodology for how policies and programs were categorised as high/low priority for evaluation
 - a plan on how/when significant policies and programs will be evaluated over the next three years
 - how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges were centred as part of the prioritisation process.

While some participants suggested that the 'ideal' would be for all policies and programs affecting Aboriginal and Torres Strait Islander people to be given 'the highest priority for evaluation' (City of Ballarat, sub. 54, p. 5), there are costs associated with evaluation, and

not all evaluations will produce meaningful or useful results. Governments also have limited resources, which means it is not practical nor feasible to conduct rigorous evaluations of all policies and programs. Priority-setting is about ensuring resources are directed to where evaluation can best inform the development of policies and programs that improve the lives of Aboriginal and Torres Strait Islander people.

This chapter:

- looks at why evaluation priorities should be set (section 6.1) and current approaches to setting evaluation priorities across and within Australian Government agencies (including for policies and programs affecting Aboriginal and Torres Strait Islander people, section 6.2)
- proposes a process and criteria for prioritising evaluations under the Indigenous Evaluation Strategy (the Strategy), including identifying an initial set of government-wide priority areas for evaluation (section 6.3)
- outlines a process for implementing the evaluation priority-setting requirements under the Strategy (section 6.4).

6.1 Why set evaluation priorities?

Setting priorities for evaluation helps ensure the policies and programs which have the most significant impact on the lives of Aboriginal and Torres Strait Islander people are evaluated. It also ensures Australian Government agencies' limited resources are directed to those policies and programs, where foreseeably, the value of accountability and/or an enhanced understanding of what works, for whom and in what context, outweighs the costs of undertaking the evaluation.

Implicit in the need to set priorities around what gets evaluated, and the resources devoted to the evaluation, is that evaluations have benefits as well as costs. In the context of benefits, the Magenta Book said 'evaluation research should only be carried out to answer questions in which there is genuine interest, and the answers to which are not already known' (HM Treasury 2011, p. 35). The Western Australian Government also said:

For some small-scale programs, it may not be cost-effective to undertake an evaluation. There may also be circumstances where necessary input from Aboriginal or Torres Strait Islander people, or sources of other essential data, cannot be obtained, rendering the evaluation infeasible or its results unclear. (sub. 74, p. 5)

Assessing the benefits of an evaluation against the costs of conducting it is not a straightforward exercise. In some cases, the benefits may not be monetised and the evidence from an evaluation can benefit other policy areas and programs. Policies and programs can also cause unintended, and therefore unexpected, harm to people's lives, and the cost of not evaluating them could be considerable.

In the context of evaluation, it is useful to think about its ‘opportunity cost’ (or how else the resources used for evaluation could be used). Resources spent on evaluation in one area will reduce the resources available for:

- policy or program delivery itself
- conducting a potentially more useful evaluation of another policy or program
- other activities across government or of those participating in evaluations (including Aboriginal and Torres Strait Islander people and organisations).

Costs of evaluation include consultant or evaluator’s fees, agency staff salaries and travel expenses, the collection of data, the time taken from those delivering the policy or program, and the time of recipients to inform the evaluation. These costs can vary considerably depending on a range of factors including, amongst other things:

- the complexity of the policy or program’s objectives, the local context and its geographical scope
- data requirements, including quality of existing administrative program data and sample size
- how the policy or program interacts with other programs
- the extent of stakeholder engagement undertaken (including users, providers and communities).

While financial costs are difficult to generalise because of these factors, they can be considerable. Based on Australian Government agency procurement contracts in 2018-19, contracts for ‘evaluation’ services ranged from \$10 000 (which is the lowest threshold for reporting on AusTender) to over \$18 million, with a median cost of about \$110 000 (Australian Government 2020a).²⁸ This is in addition to the costs these evaluations place on government and also on the communities engaged to participate in them.

Balancing the costs with the benefits of evaluation

Given these costs, it is important that the value of evaluation is greater in: informing continuous improvement in program design and implementation; ensuring accountability in policy or program effectiveness and efficiency; or adding to the knowledge base on what works, for whom and in what context.

The Productivity Commission’s *Expenditure on Children in the Northern Territory* report noted the limited value of evaluation activity when it is difficult to isolate the impacts of a particular action, particularly when there are a large number of interrelated policies and programs (that are continuously changing) affecting a community (PC 2020, pp. 278, 280).

²⁸ Tenders published in the 2018-19 financial year. Title search for tenders that contained terms beginning with ‘eval’ but excluded terms such as ‘committee member’, ‘course’, ‘training’, ‘conference’. Also excluded the category of ‘individual health screening and assessment services’.

Academics from the Centre for Aboriginal Economic Policy Research at the Australian National University also suggested that a risk of the Strategy is that it could ‘... lead to an increased number of evaluations that are focussed on marginal or minor programs to the exclusion of evaluation of the strategically significant programs’ (Tony Dreise et al., sub. 33, p. 2). They said this trend is already apparent in the current schedule of evaluations outlined in the *2018-19 Annual Evaluation Work Plan: Indigenous Advancement Strategy*, which covers ‘35 largely minor program interventions’ (sub. 33, p. 2).

This highlights the value of taking a more strategic approach to prioritising evaluations, with a focus on improving the quality of evaluation outcomes rather than simply doing more evaluations. As one participant said:

A small number of high quality comprehensive evaluations of institutionally significant policies and programs will ultimately have a much greater impact than a multitude of smaller evaluations of marginally significant programs and projects. (Michael Dillon, sub. 16, p. 2)

A priority-setting framework can also improve transparency and accountability in decision making by ensuring there is a systematic basis for determining and distributing resources for policy and program evaluation. This means program users and the broader community have oversight over the process of prioritising what gets evaluated. In the context of the Strategy, Aboriginal and Torres Strait Islander people, and the Australian community more generally, should have confidence that the resources for evaluation are used to best inform the development of policies and programs that improve the lives of Aboriginal and Torres Strait Islander people.

There is also a danger that, without a systematic basis for determining what gets evaluated, key policies and programs may fail to get evaluated. This is particularly important for ‘mainstream’ policies and programs that have significant impacts on Aboriginal and Torres Strait Islander people. The prioritisation process will ensure these policies are on the table.

A whole-of-government evaluation priority-setting process could also result in better coordination of policies and programs across government. When identifying priority policy or program areas, there will be an opportunity for agencies to collaboratively identify gaps and duplication of efforts, and ways to improve how they work together to achieve whole-of-government priority outcomes.

6.2 Approaches to setting evaluation priorities

Priority-setting of government activity is undertaken at many levels across government to ensure resources are allocated to meet government objectives. It is not clear, however, that systematic approaches to priority-setting (as exists in budgetary processes), apply across government when it comes to evaluation activity.

Some Australian Government agencies have formal processes for prioritising evaluations. While these are not specific to evaluating the impact of policies and programs on Aboriginal

and Torres Strait Islander people, the processes used provide lessons for priority-setting under the Strategy.

Limited evaluation priority-setting across government

Whole-of-government priority-setting can help governments respond to complex policy challenges that cut across agencies. To date, priority-setting across government has generally been limited to defining priority areas for policy and program *actions* rather than for the evaluation of these activities.

Arguably, the budget process is where governments are most transparent about whole-of-government priorities — and these priorities influence how funding is allocated in each budget round. Budget priorities are generally defined by senior government ministers (through the Expenditure Review Committee of Cabinet) to help inform submissions by portfolio ministers in applying for new funding proposals. However, it is not clear how these priorities are determined or that outcomes from evaluations are considered as part of the process.

For policies and programs affecting Aboriginal and Torres Strait Islander people, the Australian Government's main mechanism for outlining whole-of-government priorities is the Closing the Gap initiative. While these priorities reflect the collective commitment of all governments in Australia through the Council of Australian Governments, there is little information about how they were developed. However, what we know is that the priority-setting process under the original Closing the Gap initiative did not involve much engagement with Aboriginal and Torres Strait Islander people — a review by the Australian National Audit Office found that there was limited consultation in the development of the Closing the Gap framework (ANAO 2019a, p. 30). The review also found that policy and program outcomes linked to the Closing the Gap framework were not being effectively monitored and assessed, and an evaluation framework had not been established (ANAO 2019a, pp. 56–57).

As discussed in chapter 1, under the 2019 Partnership Agreement on Closing the Gap, a formal partnership arrangement was developed between Aboriginal and Torres Strait Islander people (through the Coalition of Peak organisations), the Australian Government, state, territory and local governments (box 1.4). However, the focus has been on a set of priority policy and/or outcomes to be pursued by governments, and while a review mechanism is proposed to assess performance of the Closing the Gap initiative, evaluation of specific policies or outcomes has not yet been prioritised.

The Australian Government previously had more formal processes for setting whole-of-government priorities on evaluations of policies and programs impacting

Aboriginal and Torres Strait Islander people (chapter 2). In the mid-2000s there were two bodies with evaluation functions that could examine cross-portfolio policy areas:

- the Office of Evaluation and Audit (Indigenous Programs) (OEA) within the then Department of Finance and Deregulation, which was a statutory function transferred in 2005, when the Aboriginal and Torres Strait Islander Commission was abolished
- the Office of Indigenous Policy Coordination, which was established in 2004, originally part of the Department of Immigration and Multicultural and Indigenous Affairs, before its functions were transferred to the Department of Families, Community Services and Indigenous Affairs (ANAO 2012, pp. 14–15).

Both these bodies published annual work plans that outlined priority evaluations. And while there is limited information on how these priorities were developed, it appears that they were inwardly focused. For example, the Office of Indigenous Policy Coordination's three-year evaluation plan only mentions consulting with other government agencies. The annual review of the plan also only talks about consulting '... Australian Government agencies working in Indigenous affairs to ensure that planned evaluation activities are targeting the areas of most need' (OIPC 2006, p. 3). The Director of Evaluation and Audit at that time also noted that one of the most challenging areas of the OEA was:

... to understand how it can create value for Indigenous people. Under the [*Aboriginal and Torres Strait Islander Act 2005* (Cwlth)], OEA has no capacity to report directly to Indigenous people and the Department's Key Performance Indicators for OEA focus on Ministerial and agency satisfaction, implementation of recommendations and timeliness. (Plowman 2008, p. 6)

More formal evaluation priority-setting processes within agencies

Some Australian Government agencies have formal processes for priority-setting of evaluation activity — although these vary across agencies. In general, those agencies that have a formal evaluation framework in place outline some proportional approach to deciding what (and to what extent) a policy or program is evaluated within each agency (box 6.1).

Agencies typically use a principle-based approach to prioritising evaluation activity based on the characteristics of a given activity, such as the size, significance or risk profile of the policy or program. Based on this assessment, agencies then decide the level of evaluation effort for a given policy or program. Some agencies provide further guidance on how this assessment should be made, such as through a formal prioritisation tool, and also establish certain requirements based on the priority level a policy or program is given. For example, under the former Department of the Environment's *Evaluation Policy 2015–2020*, interventions classified as 'significant' by a program area were required to:

- Consult the Evaluation team in [the Policy Analysis and Implementation Division] prior to finalising the evaluation brief (or terms of reference).
- Ensure the team undertaking the evaluation and the steering committee [include] independent members.

-
- [Consult] with community and business groups affected by the intervention ... [as] part of the evaluation design.
 - Provide a copy of the evaluation report to the Evaluation team in [the Policy Analysis and Implementation Division]. (Department of the Environment 2015, p. 10)

... but transparency of these processes vary

While the priority-setting principles are broadly consistent across agencies, the governance arrangements differ (including around transparency, the formality of approach to setting of priorities, and requirements established for evaluations once priorities have been set). For example, the former Department of Industry, Innovation and Science, the Department of Foreign Affairs and Trade, and the National Indigenous Australians Agency have centralised evaluation functions which coordinate priority-setting of key evaluation activities in their respective agencies, and manage and report on progress against these through a forward work plan (reviewed annually) — the Department of Industry, Innovation and Science, unlike the others, did not publish the plan in full (DIIS 2019a, p. 15).

While aspects of evaluation planning may feature in some agencies' public reporting, including through their obligations to report on performance under the *Public Governance, Performance and Accountability Act 2013* (Cwlth), this reporting seldom covers how assessments are made on what to evaluate. Importantly, it is unclear whether the community is consulted during the priority-setting process, and what avenue they have to provide feedback on the evaluations each agency has selected. This is the case even for those agencies that publish an annual list of planned evaluations.

Box 6.1 **Examples of priority-setting within government agencies**

Department of Industry, Innovation and Science

The former Department of Industry, Innovation and Science's *Evaluation Strategy 2017–2021* outlines how evaluation effort is prioritised for each of its programs. Evaluation effort is scaled, using a tiered approach, based on the value, impact and risk profile of a program, assessed using the following criteria:

- total funding allocated for the program
- internal priority (importance to the Department's and Australian Government's goals)
- external priority (importance to external stakeholders)
- overall risk rating of the program
- track record (previous evaluation, the strength of performance monitoring and lessons learnt).

Based on this criteria, programs are ranked into one of three tiers. Programs assessed as Tier One are marked as evaluations of highest priority and strategic importance, whereas programs in Tier Three are lower priorities for evaluation.

The Department also uses the tiering system to identify the evaluations of highest priority and strategic importance, which are then included in the Department's four-year evaluation plan. The plan is maintained by a central unit, who reports progress to the Department's Executive. Elements of the plan are published in corporate performance reporting documents.

Department of Foreign Affairs and Trade

The Department's *Aid Evaluation Policy* outlines how evaluations are prioritised at two levels across the Department. At the program level, each program area undertakes an annual process to identify and prioritise a reasonable number of evaluations. While there is a minimum set of evaluations required to be undertaken by each program (depending on program size), programs have the flexibility to initiate and manage what gets evaluated.

A prioritisation tool is provided to aid program areas determine the highest priority issues their evaluations should focus on. This includes areas where:

- there are significant evidence gaps
- risks threaten the effectiveness of the Department's work
- there is a heightened need for accountability
- budgets are large
- investments are high profile or otherwise important for the Australian Government or their partners.

A central unit within the Department, the Office of Development Effectiveness (ODE), can also independently initiate and manage a set of 'strategic evaluations', which focus on broad aid policy directions or specific development themes.

(continued next page)

Box 6.1 (continued)

The selection of topics is guided by the following principles:

- policy relevance: assessing progress on implementing Australian Government policy priorities
- potential learning benefits: focusing on where there is a strong demand for insight, information or guidance
- coverage and materiality: over time ODE's evaluations should provide appropriate coverage of expenditure by sector, theme, country and region
- risks to aid effectiveness: examination of issues which pose a challenge to the Australian aid program and its reputation.

Each year ODE compiles the Department's Annual Evaluation Plan. The plan, which is endorsed by the Independent Evaluation Committee and approved by the Secretary prior to being published on the ODE's website, outlines both the strategic and program evaluations to be conducted over the calendar year.

National Indigenous Australians Agency (NIAA)

In 2018, the Department of the Prime Minister and Cabinet released an Indigenous Advancement Strategy evaluation framework, which described high-level principles for how evaluations of Indigenous Advancement Strategy programs should be conducted, and outlined future capacity-building activities and broad governance arrangements. As part of this process, the NIAA has published an annual work plan of evaluations. To develop these work plans, the NIAA catalogued existing evaluation activities, and conducted a 'bottom-up' process to identify new activities. According to the Australian National Audit Office, this involved three steps:

- the evaluation branch of the NIAA consulting other areas of the agency to identify and develop evaluation proposals
- the evaluation branch assessing the technical merit, significance, contribution and risk for submitted proposals
- oversight bodies, such as the Indigenous Evaluation Committee and the Executive Board of the NIAA considering and endorsing the draft plan.

Prioritisation under this framework considers significance, contribution and risk. According to the framework, 'significant, high risk programs/activities will be subject to comprehensive independent evaluation.'

Department of the Environment

The former Department of the Environment prioritised evaluation by taking into account scale, as well as other program characteristics including size, significance and risk profile. To help determine evaluation priorities, the Department used a tiered ranking system which characterises programs into one of three categories (low, medium, or high risk). Each category has a range of program characteristics, and is linked to corresponding likely characteristics of evaluation. For example, a program determined to be high risk would require a formal evaluation process, a high resource allocation, extensive consultation and a wide public release. Whereas, a program determined to have characteristics of low risk, would require an informal evaluation process, low resource allocation, and be completed internally.

Sources: ANAO (2019b, pp. 7, 33, 38); DIIS (2019a); DFAT (2017a, 2017c, p. 4); Department of the Environment (2015); DPMC (2018).

6.3 Priority-setting under the Indigenous Evaluation Strategy

Establishing government-wide evaluation priorities

Because the Strategy aims to improve the lives of Aboriginal and Torres Strait Islander people, evaluation priorities should focus on informing better policy outcomes in areas that contribute most significantly to this objective. Given this relationship, the setting of *evaluation priorities* should largely reflect *priority policy outcomes* that the Australian Government and the Indigenous community identify as being important to improving the lives of Aboriginal and Torres Strait Islander people.

Policy priority-setting processes (such as budget processes and whole-of-government policy reforms, agendas or strategies) could inform priorities for evaluation; however, because the overarching principle of the Strategy is to centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (chapter 11), the priority-setting process should:

- reflect Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges
- have a clear program logic on how the policy priority is expected to improve the lives of Aboriginal and Torres Strait Islander people.

Policy priorities from Closing the Gap

Policy priorities identified under the Joint Council on Closing the Gap are a good starting point to inform government-wide evaluation priorities. The Partnership Agreement on Closing the Gap (the partnership is between all levels of government and peak Aboriginal and Torres Strait Islander organisations) prioritises efforts to drive intergenerational change and increase opportunities for Aboriginal and Torres Strait Islander peoples to reach their full potential by leading the lives they value. The *draft* framework has a strong focus on early investment and a life course approach and recognises that safe environments, health and wellbeing and education have profound impacts on an individual's socio-economic outcomes later in life (NIAA, sub. 86, p. 4).

Aligning priorities with the Partnership Agreement on Closing the Gap process was endorsed by several participants, including Empowered Communities (sub. 41, p. 11), the Royal Australasian College of Physicians (sub. 77, p. 6), the Tasmanian Government (sub. 100, p. 5) and the Western Australian Government (sub. 74, p. 10). For example, the Western Australian Government said:

The Strategy should align with the evaluation requirements for the Closing the Gap Refresh, which will establish the primary criterion for determining evaluation priorities under the Strategy. This will ensure that ... Indigenous perspectives of priority-setting will be heard and incorporated through the Closing the Gap consultation and accountability architecture ... (sub. 74, p. 10)

The NIAA suggested that the strategic alignment of the Partnership Agreement on Closing the Gap targets and the Strategy ‘... needs to occur prior to deciding on evaluation topics and methods’ (sub. 86, p. 4).

Defining evaluation priorities from policy priorities

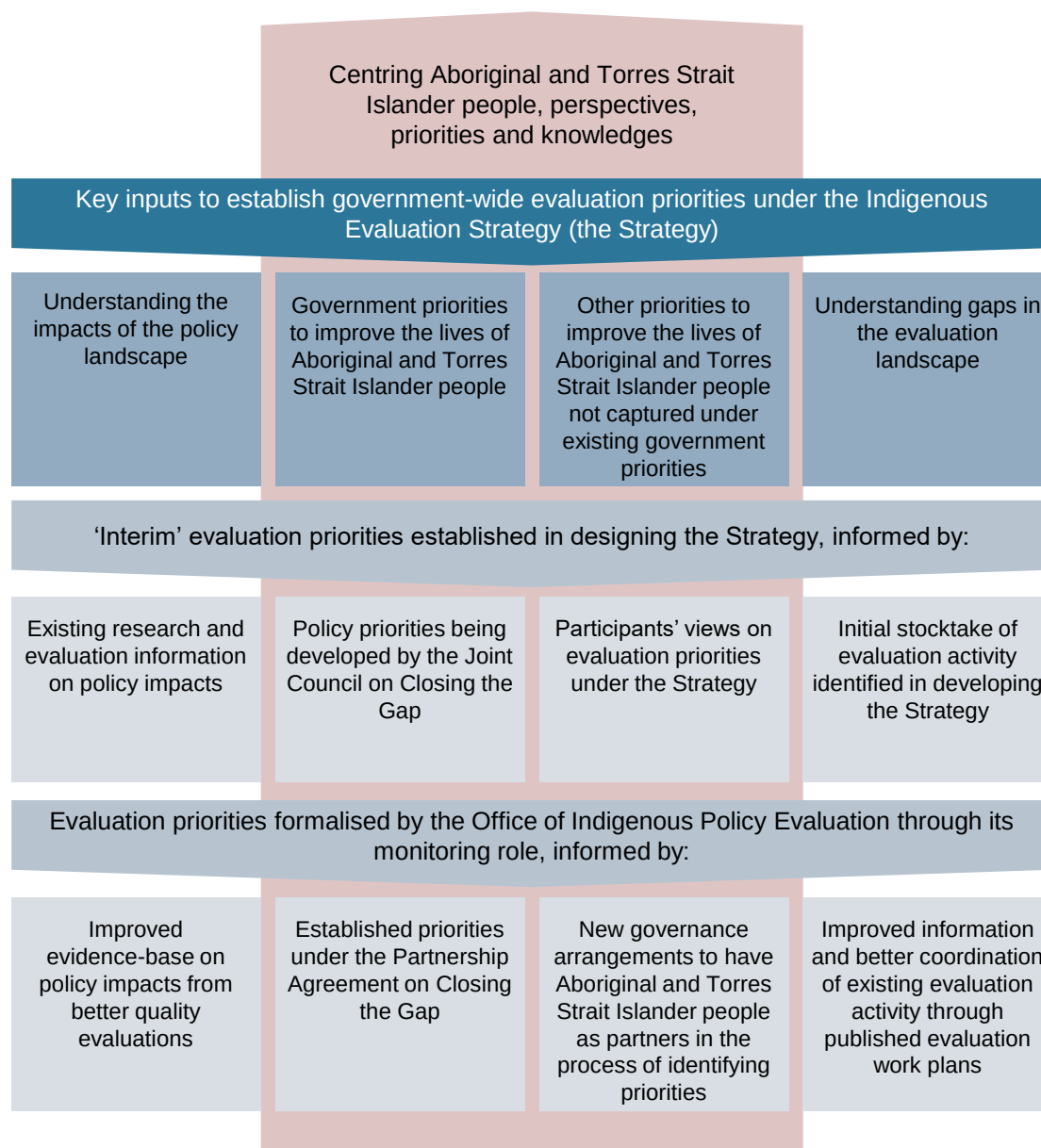
The Partnership Agreement on Closing the Gap is currently not focused on evaluation of specific policies and programs. The outcomes therefore require some refinement to help establish a set of government-wide evaluation priorities (figure 6.1). This includes considering policy or program priorities aimed at improving the lives of Aboriginal and Torres Strait Islander people that are not captured through the Closing the Gap initiative, as well as the state of the evidence base in these priority areas.

Understanding where improving the evidence base will add value involves looking at the:

- scope of government activity in these priority areas, such as, expenditure, the number, size and scope of programs, and the prominence of these programs in the government’s agenda etc.
- potential impact of activities (policies or programs) in these priority areas
- current evaluation activity (and importantly the gaps) in these priority areas.

The task of establishing government-wide evaluation priorities should initially sit with the proposed independent Office of Indigenous Policy Evaluation (OIPE), and then the proposed Centre of Evaluation Excellence when it is established. In line with the overarching principle of the Strategy, the Commission is also proposing an Indigenous Evaluation Council, which will help to ensure the evaluation priorities are established in partnership with Aboriginal and Torres Strait Islander people (chapter 7).

Figure 6.1 Process for formalising government-wide evaluation priorities



An interim set of government-wide evaluation priorities

While the Commission has been asked to identify evaluation priorities, this is not possible without having the governance arrangements in place to establish a partnership arrangement with Aboriginal and Torres Strait Islander people. Developing evaluation priorities without partnering with Aboriginal and Torres Strait Islander people is inconsistent with the overarching principle of the Strategy of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges.

The Commission has, however, identified a broad ‘interim’ set of evaluation priorities based on the priorities identified by the Council of Australian Governments (COAG) and the Joint Council on Closing the Gap, as well as:

- what Aboriginal and Torres Strait Islander people and communities told us on this project about what the priorities should be
- an initial assessment of the existing evidence base.

The Commission plans on undertaking more targeted consultation with Aboriginal and Torres Strait Islander people on what the evaluation priorities should be between the draft and final report.

While there is some evidence of the impacts of existing policies and programs on Aboriginal and Torres Strait Islander people, there are still many gaps in the evidence-base across government (chapter 3). This makes it difficult at this stage to make judgements on how one policy area should be prioritised over another. For example, the Commission’s report on *Expenditure on Children in the Northern Territory* found several examples of evaluations undertaken on tertiary intensive family support services, but more limited examples of evaluations on preventative services aimed at keeping children safe and culturally strong (PC 2020). This gap in evidence-base would suggest the latter may be considered a priority for evaluation, but whether this is a priority over housing support or broader welfare support services is more difficult to determine given current gaps in the overall evidence base. It is expected that these information gaps will reduce over time, allowing better refinement of these priority evaluation areas (section 6.4).

The draft Closing the Gap priorities (as of March 2020) covered seven priority policy areas, a generic cross-system priority, and up to four priority reform areas (table 6.1). The list captures some of the specific policies and programs identified as priorities for evaluation in submissions to this project, as well as key policy objectives under the Indigenous Advancement Strategy (table 6.2). The priority areas also reflect some of the key expenditure areas identified in the *Report on Government Services* (with the omission of social security and welfare support programs) (chapter 3). This suggests that they are a useful starting point for an interim set of priority areas for evaluation under the Strategy. Reflecting their more systematic focus, the cross-system priority and four priority reform areas could be a focus for government to undertake more collaborative evaluations across several portfolios.

Table 6.1 Proposed interim government-wide evaluation priority areas
Based on draft priorities established by COAG 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)

Sources: Coalition of Aboriginal and Torres Strait Islander Peak Organisations (2018, 2020); JCOCTG (2019b).

The focus on the draft Closing the Gap priority policy areas rather than targets reflects the broader aim of the Strategy to improve the lives of Aboriginal and Torres Strait Islander people, rather than focus on the more narrowly defined indicators of 'closing the gap'. As the National Justice Project noted:

We believe that there are important lessons to be learned from the framing of Closing the Gap 'targets' which confuse outcomes and indicators. These targets purportedly track progress towards 'closing the gap' but do not address the systemic issues that underlie important impacts and outcomes for Aboriginal and Torres Strait Islander peoples. (sub. 51, p. 4)

And according to the Coalition of Peaks involved in the Joint Council on Closing the Gap, there is broader support around the policy priority outcomes (relative to the targets) amongst Aboriginal and Torres Strait Islander people.

At the engagements, most Aboriginal and Torres Strait Islander people supported the priority areas (such as health, education, housing, justice, children and families, and economic development) for which targets have been developed. The feedback on the targets themselves, including the wording, outcomes, measures and focus, was often that they needed to be changed, with suggestions offered.

Of survey respondents, 66% supported the draft Closing the Gap targets agreed by COAG. (Coalition of Aboriginal and Torres Strait Islander Peak Organisations 2020, p. 7)

INFORMATION REQUEST 6.1

The Commission proposes that the interim evaluation priorities for the Indigenous Evaluation Strategy be broad and based on the Council of Australian Governments and the Joint Council on Closing the Gap draft policy priorities. Is this appropriate?

If so, are there any priority areas missing from those currently identified through the Council of Australian Governments and the Joint Council on Closing the Gap?

If not, what specific policy or program areas should be the focus of better quality evaluations?

Prioritisation process within agencies

Identifying scope of policies and programs for evaluation: Flows and Stock

The set of government-wide evaluation priority areas (table 6.1) can help agencies identify the policies and programs within their remit that are considered to be in scope of the Strategy. The prioritisation process is about knowing what policies and programs are in place across Australian Government agencies that promote these government-wide priority areas, and that should be considered for evaluation under the process illustrated in figure 6.2. (Table 6.2 illustrates how government-wide evaluation priorities can capture specific policies and programs identified from submissions to this project.)

All government departments, as stewards of policies and programs within their portfolio responsibilities, should at a minimum identify the policies and programs that contribute to government-wide evaluation priorities — along with others in their remit that aim to improve the lives of Aboriginal and Torres Strait Islander people. This should cover both new policy and program proposals as well as existing policies and programs (draft Strategy action 2).

It should also apply to both policies and programs directly targeted at Aboriginal and Torres Strait Islander people, as well as mainstream policies and programs that have a significant impact on the lives of Aboriginal and Torres Strait Islander people. The prioritisation process will be important for identifying which mainstream policies and programs have the largest impact on the lives of Aboriginal and Torres Strait Islander people, and therefore where the evaluation focus should be for the purpose of this Strategy.

Table 6.2 Specific evaluation priority areas identified in submissions^a

Priority 1: Families, children and youth^b

- Early childhood — the first 1000 days of a child's life (Maggie Walter, sub. 112, p. 2)

Priority 2: Health

- Suicide prevention (Generation One, Minderoo Foundation, sub. 5, p. 5)
- Indigenous Australians' Health Programme (AMSANT, sub. 81, p. 5; Royal Australasian College of Physicians, sub 77, p. 6)
- Hearing health (Audiology Australia, sub. 27, p. 2)
- Mental health (IUIH, sub. 53, p. 12)
- Introduction of: Framework Agreements in Aboriginal health, and State and Territory-based planning forums (from the late 1990s); section 100 of the *National Health Act 1953* (Cwlth) on access to Pharmaceutical Benefits Scheme medicines in remote areas since July 1998; and access to Medicare Benefits Schedule billing in July 1996 for services to Aboriginal communities (Central Australian Aboriginal Congress, sub. 48, p. 7)
- The impact of: having funding and administrative responsibility for Aboriginal health within the Department of Health since 1995 compared with its place as part of a general Aboriginal Affairs portfolio; and of whole-of-population public health strategies that benefit Aboriginal people such as population-level alcohol supply reduction measures and tobacco control measures (Central Australian Aboriginal Congress, sub. 48, pp. 7–8)
- National Aboriginal and Torres Strait Islander Health Plan (AMSANT, sub. 81, p. 5; Central Australian Aboriginal Congress, sub. 48, p. 8)
- Mainstream programs such as the NDIS/aged care for Aboriginal people (AMSANT, sub. 81, p. 6)

Priority 3: Education^b

- Indigenous higher education policies and programs (Clair Anderson, sub. 105, p. 5)

Priority 4: Economic Development^b

- First Nations economic development, labour market policy and wealth (including the Community Development Programme and the Indigenous Procurement Policy) (Dreise et al., sub. 33, p. 2)

Priority 6: Justice, including youth justice

- Discrimination in the delivery of government programs and services, and in interactions with the justice system (National Justice Project, sub. 51, p. 4)

Priority 7: Land and waters

- Administration of the *Native Title Act 1993* (Cwlth) (Dreise et al., sub. 33, p. 2)

Cross system priority

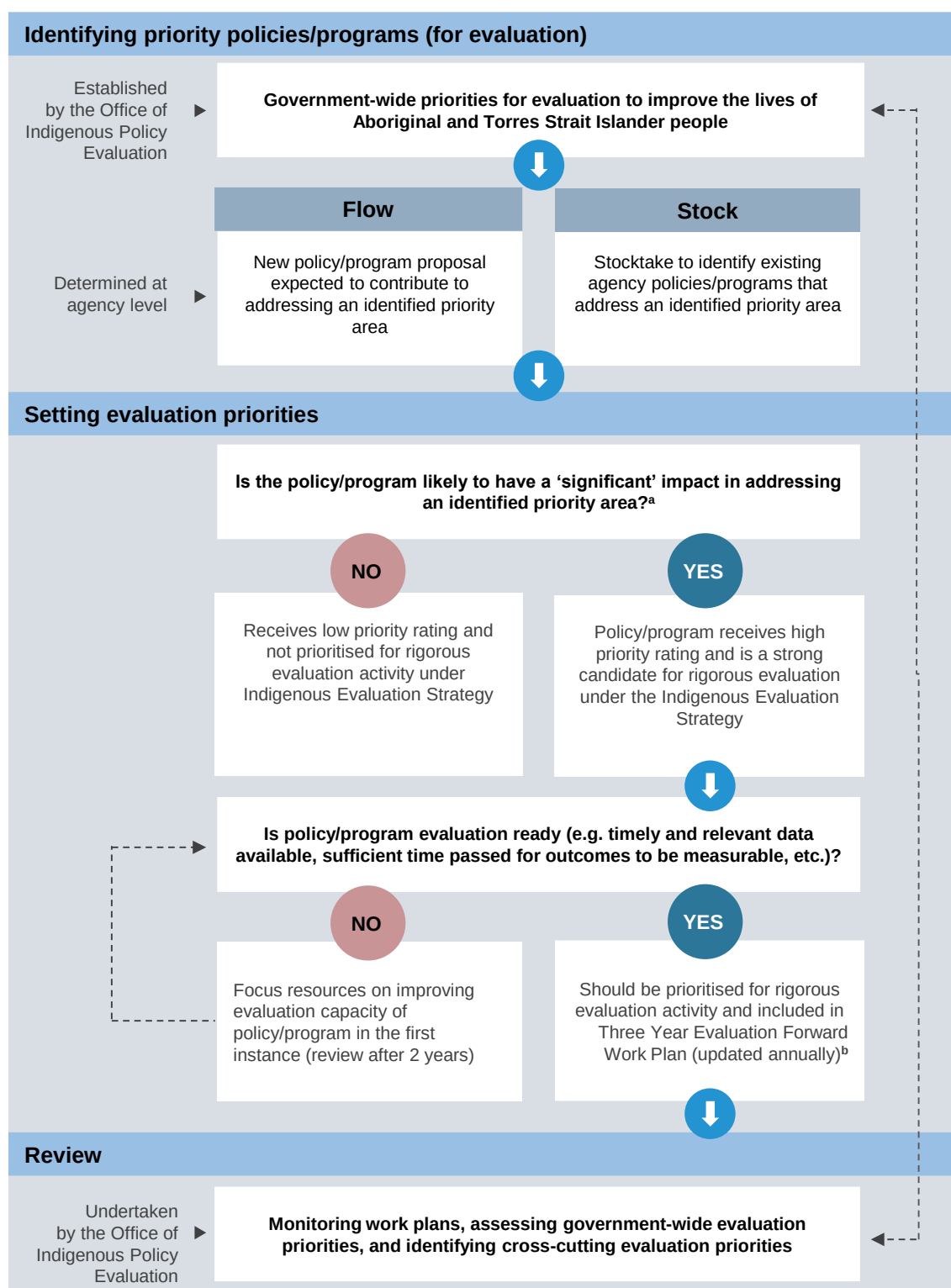
- The need for allocating specific infrastructure funding to support enhanced service accessibility in urban settings, including expanded clinic development (IUIH, sub. 53, pp. 11–12)
- Digital inclusion and digital literacy (First Nations Media Australia, sub. 30, p. 11)
- Implementation of Royal Commissions, reviews and inquiries (AMSANT, sub. 81, p. 5), for example: Royal Commission into Aboriginal Deaths in Custody (1991); Bringing Them Home report (1997); Little Children are Sacred (2007); Royal Commission into the Detention and Protection of Children in the Northern Territory (2017) (Central Australian Aboriginal Congress, sub. 48, p. 8)
- Implementation of: the Indigenous Advancement Strategy; and policies related to the social determinants of health (e.g. remote housing and the Community Development Programme) (AMSANT, sub. 81, p. 5)

Priority reform areas

- Effectiveness of/access to community controlled sector (AMSANT, sub. 81, p. 4; Central Australian Aboriginal Congress, sub. 48, p. 2; National Mental Health Commission, sub. 19, p. 4)

^a No specific areas identified under priority area 5 (Housing). ^b Refers to areas that relate to focus areas under the Indigenous Advancement Strategy: Addressing domestic and family violence (healthy and safe home and community); Children going to school (education); Conditions and incentives for Indigenous Australians to participate in the economy and broader society (employment, economic development and social participation).

Figure 6.2 Priority setting under the Indigenous Evaluation Strategy



^a 'Significant' as defined by the priority-setting criteria outlined in this chapter. ^b Evaluation priorities may also be set outside the Indigenous Evaluation Strategy, such as through sunset clauses, reviews of major programs determined when programs are introduced, etc. These could also be included in the rolling Three Year Evaluation Forward Work Plans.

With new policies and programs, agencies would be required to identify if the activity proposed for Cabinet approval, under the New Policy Proposal process, will contribute to any of the government-wide evaluation priorities. If so, it should be considered in scope for evaluative activity under the Strategy. This will complement the proposed requirement on agencies to embed evaluation planning as part of the New Policy Proposal stage, and to ensure that evaluation is considered and resourced early in the policy development process for any policy that is expected to impact the lives of Aboriginal and Torres Strait Islander people (chapter 7; draft Strategy action 2).

To deal with the stock of policies and programs, agencies should undertake a stocktake of existing policies and programs to be considered in scope of the Strategy. The Commission has not specified a process for undertaking a stocktake — agencies should decide on an approach that best suits them based on the potential reach of the Strategy within their own portfolio agencies. Potential approaches for a stocktake could include identifying key policies and programs through:

- management approaches — such as targets to evaluate a certain proportion of existing policies and programs under the Strategy in a given year
- programmed reviews — identifying relevant policies and programs as part of scheduled reviews into particular policy or program area, for example, through sunseting reviews or other reviews embedded in existing legislation or post-implementation reviews
- ad hoc reviews — including in-depth review of the potential impact of policy or program areas on Aboriginal and Torres Strait Islander people.

Regardless of the approach taken, the stocktake should be transparent, including publishing details of:

- the methodology used to undertake the stocktake (this could include outlining the criteria used to assess what policies and programs are in scope). The process could also involve the agency setting specific targets for evaluation activity or a whole-scale review of key legislations and administrative functions
- the governance arrangements around the stocktake, including who provided oversight, and who was involved in doing the stocktake
- how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges were centred in the stocktake process
- the outcomes of what policies and programs were found to be in scope of the stocktake.

These details would be outlined in a rolling Three Year Evaluation Forward Work Plan (discussed below).

Deciding what to evaluate and how

Once policy and program priorities are clearly defined, the next stage is to determine which of these should be evaluated, and how.

There is a fairly standard set of criteria used for deciding what to evaluate. These are summarised in table 6.3. While participants had different views on what criteria should be used to determine key priorities (box 6.2), most suggested criteria commonly used in priority-setting frameworks.

Table 6.3 Criteria for priority-setting

<i>Criterion</i>	<i>Summary</i>
Impact of program	This criterion is based on how likely the program will impact the lives of Aboriginal and Torres Strait Islander people. It could take account of the total Aboriginal and Torres Strait Islander population affected or whether a policy or program disproportionately affects Aboriginal and Torres Strait Islander people.
Strategic significance	This principle is based on 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	This principle is based on the program's 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 rights and lives of Aboriginal and Torres Strait Islander people.
Program expenditure	This principle is based on overall funding directed to 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.

In the context of priority-setting of evaluations for policies and programs affecting Aboriginal and Torres Strait Islander people, underpinning each criterion should be the centring of Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. For example, when considering the risk profile of an education program in scope of the Strategy, the risks as they relate to Aboriginal and Torres Strait Islander people should be assessed (such as the risk of a program with limited evidence of working in Aboriginal and Torres Strait Islander communities). This highlights the importance of establishing effective and enduring partnerships with Aboriginal and Torres Strait Islander people when identifying policy and program priorities, as they will also be important for determining evaluation priorities and assessments of impact.

Agencies should decide on the relative importance of each criterion (and how it should be assessed) within their own portfolio. Agencies with smaller budgets may set expenditure thresholds lower relative to departments with considerably larger programs or budgets. In a given year, an agency may implement a large number of pilots and therefore risk may be weighted relatively higher than other criterion during that year's assessment period.

Box 6.2 **Participants' views on criteria for priority-setting**

The priority-setting criteria for evaluation proposed by participants included:

- The history of evaluation for each program:
... reflecting on the frequency of evaluations, prior evaluations and significance of the findings/recommendations and the acuteness or severity of the issues which the program seeks to ameliorate. Programs which maintain a high level of community dissatisfaction, suboptimal outcomes or which have a poor record of evaluation should be prioritised. (Jobs Australia, sub. 57, p. 9)
- The priorities of Aboriginal and Torres Strait Islander people and organisations (the Western Australian Government, sub. 74, p. 10; Social Ventures Australia, sub. 83, p. 3).
- The scale of the program in terms of the number of Aboriginal and Torres Strait Islander people impacted, and scale of public expenditure (Sophia Couzos, sub. 92, p. 2; the Western Australian Government, sub. 74, p. 10; Social Ventures Australia, sub. 83, p. 3; VACCA, sub. 26, p. 9).
- Programs with 'limited to no evidence underpinning their approach' (Social Ventures Australia, sub. 83, p. 3), including new and pilot programs (Sophia Couzos, sub. 92, p. 2; the Western Australian Government, sub. 74, p. 10). According to the Western Australian Government, 'it is highly recommended that pilot programs are evaluated to assess their feasibility and potential scalability' (sub. 74, p. 10).
- Programs of particular state and territory or national significance (Sophia Couzos, sub. 92, p. 2) and that align strategically with government objectives (the Western Australian Government, sub. 74, p. 10).
- Other factors, including, 'whether the program is being considered for discontinuation ... [program] complexity ... and degree of risk (e.g. to Aboriginal communities and Government)' (the Western Australian Government, sub. 74, p. 10).

Agencies should develop a fit-for-purpose multi-criteria assessment tool, based on the criteria identified above to help categorise the relative significance of all its policies and programs in scope of the Strategy (draft Strategy action 1). For example, agencies may decide to give equal weighting to all the criteria and assess each policy or program on a five-point scale. These assessments would be done consistently by agencies across all new policies and programs (and those identified as part of the stocktake) so the aggregate score would determine their place along a continuum of overall 'significance'. Based on resources and the level of activity, agencies would then establish a threshold so that all 'significant' policies and programs will be rigorously evaluated (draft Guide, section 2, box 4).

Having multiple criteria requires government agencies to weigh up all criteria to make their decision on what policies and programs should be prioritised for rigorous evaluation. For example, a pilot intensive family support program, trialled in a small community may only have a relatively 'small' aggregate impact in terms of expenditure and population, but if the potential risk of the program to a child's wellbeing is significant, it could be prioritised for evaluation. Alternatively, an established program with significant expenditure but that has been recently evaluated, may be less of a priority for formal evaluation activity in the short term.

To ensure agencies use these criteria effectively, there are a few requirements that they should follow:

1. assessments should centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges, which requires agencies to have appropriate governance to ensure Aboriginal and Torres Strait Islander people are engaged in making these determinations (box 6.3)
2. the assessment methodology should be publicly available (and included as part of each their rolling Three Year Evaluation Forward Work Plan) so stakeholders can be clear about how decisions have been made on what has been prioritised for evaluation
3. the assessment methodology should be reviewed and updated over time to ensure it remains relevant (section 6.4).

Box 6.3 Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges at the centre of the prioritisation process

To improve outcomes for Aboriginal and Torres Strait Islander people, it is important that they define what these outcomes are and, equally, the priority they are given by government. There are few mechanisms in place for Aboriginal and Torres Strait Islander people to determine and decide on policy priorities. Constant policy changes have meant that mechanisms for Aboriginal and Torres Strait Islander people to make decisions about, and inform, policies and programs that affect their lives have been largely short-lived or absent altogether. This sentiment was captured by The University of Queensland:

It is well documented that the impact of colonisation and subsequent Government policies have resulted in a lack of Indigenous agency or voice in policy and program development. Scholars have for decades identified the impact of exclusion of Indigenous peoples in the conceptualising, development, research and evaluation of Indigenous policy and programs. (sub. 20, p. 5)

A number of participants suggested that Aboriginal and Torres Strait Islander people should be part of the process to review evaluation priorities (Cape York Institute, sub. 69, p. 9; National Mental Health Commission, sub. 19, p. 5; PwCIC, sub. 58, p. 6; VACCA, sub. 26, p. 9). For example the National Mental Health Commission said:

In line with the principles of co-design and co-production, the [NMHC] considers that evaluation priorities should be driven by Aboriginal and Torres Strait Islander peoples themselves. As the stakeholders that have the most to gain (or lose) from the outcomes of evaluations under the Indigenous Evaluation Strategy, Aboriginal and Torres Strait Islander peoples should have a central role in determining evaluation priorities. (sub. 19, p. 5)

Lower priority policies and programs

While it is important that resources are given to rigorously evaluating the most significant policies and programs, this does not mean that all other policies and programs should not be developing their own evidence base. Agencies should develop thresholds to further categorise lower levels of evaluative activity for those policies and programs which are not categorised as ‘significant’.

So while the ‘significant’ policies and programs should be adequately resourced for evaluation, those with a lower priority could focus on improving data collection on more meaningful outcome measures, monitoring and performance audits, or other evaluative activities that facilitates learning by doing (such as user surveys, reviews etc.). Maintaining a focus on these will ensure that evaluation culture is embedded in policy design and implementation across government, and also ensure that these programs are ready for more rigorous evaluation activity in the future, if required.

There may also be other evaluations that are likely set outside the Strategy, such as through sunset clauses and reviews of major programs determined when programs are introduced. For full transparency, these evaluations should also be included in the rolling Three Year Evaluation Forward Work Plan.

Evaluation readiness

While all significant policies and programs should be subject to rigorous evaluation, the timing of this evaluation should be flexible to ensure greatest value is obtained from the process. The final stage of the prioritisation process is to establish the best form of evaluation and most appropriate timing for it to take place. This should take into account:

- existing data limitations in current policy and program design that may require resources to collect better information
- time for outcomes to be reasonably measurable
- the policy environment, for example several policy changes affecting the same population group at the same time have taken place which may make attribution difficult.

Some important considerations of evaluation readiness are included in the tool developed by the former Department of Industry, Innovation and Science (box 6.4).

These issues are a constant challenge faced in evaluation, particularly in an environment where evaluative culture is not yet fully mature. It is also important to recognise that evaluations can often take time to undertake, and are most effectively done when they have been planned from the inception of the policy design process.

There are also different types of evaluation that are more suited to different stages of a policy or program’s implementation (chapter 4) and agencies may need to put in place enabling activities to support future evaluations. For example, in their forward work plan the National Indigenous Australians Agency made note of a number of planned enabling activities for those policies and programs that are not yet ready for full evaluation. These activities include evaluation strategies, capability development projects, and data improvement projects (NIAA 2019, pp. 12–17).

In the early days of the Strategy’s implementation, it may not be feasible to subject all significant policies and programs to rigorous evaluations. This is not to suggest that no

activity should be initiated for these policies and programs, rather than agencies should have a plan to get them ready for evaluation in a reasonable timeframe.

Box 6.4 **The Department of Industry, Innovation and Science's Evaluation Ready process**

Many Australian Government departments' evaluation strategies articulate some form of process that forces the early consideration of evaluation in a program's life. The value of such processes however goes beyond the mere smoothing of the evaluation process when it comes to be taken up at a later stage, rather it helps sharpen the objective and design of the program while still in the design phase. The *Evaluation Ready* process outlined in the former Department of Industry, Innovation and Science's *2017–2021 Evaluation Strategy* provides a good example of early consideration of evaluation that takes the form of five steps.

1. An **initial meeting** between the relevant policy and program delivery areas, and representatives of the Department's in-house evaluation unit to cover relevant information including background information and possible outcomes.
2. A **program logic model**²⁹ is developed by the evaluation unit using information from the initial meeting, with feedback from policy and delivery areas.
3. A **data matrix** is developed, helping formulate appropriate evaluation questions and areas of inquiry. A data matrix will typically look to identify success indicators, their data sources, the responsible party, analysis methods and how the data should be interpreted.
4. From here, the **evaluation strategy** sets out proposed timing, resourcing, methodologies, risks and responsibilities and is signed off by the policy General Manager. The program's evaluation strategy documents the thinking behind the evaluation's approach and becomes a resource for future planning.
5. The Manager of the evaluation unit endorses the program as 'Evaluation Ready' and the package is presented to the Program Assurance Committee.

Source: DIIS (2019a).

Three Year Evaluation Forward Work Plans

The process of evaluation prioritisation should be transparent (this aligns with the transparency principle of the Strategy).

The Commission recommends (draft Strategy action 1) that each department centrally coordinate, 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

²⁹ A program logic model is a visual representation of a policy or program's theory of change, typically showing inputs, activities and the resulting intermediate and ultimate outcomes the policy or program is aiming to achieve.

-
- the methodology for how the agency categorised the high priority policies and programs based on the criteria identified above
 - 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.

The annual work plan would include new policies and programs implemented over the year, as well as an appropriate number of those identified as part of the stocktake of existing policies and programs. One of the questions for the proposed OIPE in reviewing the work plans would be: has the relevant agency identified a sufficient proportion of their existing set of policies and programs?

The plan should be publicly released on the agency's website, as well as provided to the OIPE so it can review the agency's performance against the Strategy (chapter 10). There should also be mechanisms for the community to provide feedback on the adequacy of these plans, which should also be made public. This would further raise the incentives for departments to ensure the quality of their prioritisation process.

Departments would make appropriate judgements on whether larger agencies within their portfolio should publish their own individual forward work plans. This assessment would need to balance the relative burden on agencies to undertake this task, against the value for agencies to build their own central evaluation capability. Centralising the reporting of evaluative activity to a department may also facilitate greater coordination of evaluative effort and resources across common policy areas.

6.4 Initial implementation and a process of review

Timelines of stocktake and rolling Three Year Evaluation Forward Work Plans

Within the first six months of the Strategy being implemented, departments and major agencies should have:

- adopted a multi-criteria tool to assess the significance of policies and programs that are in scope of the Strategy
- made changes to New Policy Proposal processes to assess the significance and evaluability of new policies and programs that impact Aboriginal and Torres Strait Islander people.

Within a year, all departments and major agencies should have developed an initial rolling Three Year Evaluation Forward Work Plan. The initial plan is unlikely to include a full stocktake of existing policies and programs but should include the evaluation of mainstream policies and programs that contribute to the 'interim' set of priority evaluation areas.

Interaction with existing processes

As a principle-based, and largely devolved framework, the priority-setting process should align with existing processes within agencies. Agencies should not be required to adopt all assessment criteria outlined in the Strategy if it has made an assessment that its existing criteria adequately covers the main principles outlined in this chapter.

There should also be some flexibility in how evaluation priorities within the Strategy coexist with other evaluation priorities identified by agencies. In some cases, there may be overlap in the scope of policies and programs that are covered — this could particularly be the case for mainstream policies and programs.

A process of reviewing evaluation priorities

As discussed in chapter 10, the OIPE would review agencies' rolling Three Year Evaluation Forward Work Plans and evaluation activity to:

- assess whether they have complied with requirements set out in the Strategy — and the quality of their compliance activity
- make strategic assessments of cross-portfolio policy areas where agencies could collaborate on to undertake scheduled evaluative work
- formalise a set of evaluation priorities focused on improving the lives of Aboriginal and Torres Strait Islander people.

7 Embedding a culture of evaluation and learning

Key points

- The Indigenous Evaluation Strategy (the Strategy) will be most effective in an environment where evaluation is valued for accountability, learning and evidence-based decision making. This is where agencies want to know how their policies and programs are performing, are prepared to experiment (in areas of policy where answers are not known), share learnings, and use evaluation results in policy making.
- While there is strong agreement across Australian Government agencies that policy should be informed about what works, in practice (in part because of the realities of policy making) evaluation findings are often not used in policy decision making or to inform agencies about where to invest taxpayer funds.
- Evaluations are more likely to be used by policy makers if they address the questions that decision makers want answered, and if evaluators collaborate and build relationships with those affected by the policy or program – in this context, Aboriginal and Torres Strait Islander people, communities and organisations. This points to the importance of evaluation being embedded in policy planning and policy making processes.
- Policy makers need to know about evaluation evidence (evaluation results need to be published) to use it. Transparency of evaluation results can also improve the quality of evaluations, facilitate more informed public debate about the effectiveness of policies, and increase pressure for follow-up on evaluation findings.
- Good knowledge management, which is missing under current arrangements, is also key to making evaluation evidence easy for program managers and policy makers to use. Evaluation results are more likely to be used if they are synthesised (helping policy makers access large bodies of evidence) and tailored for specific users.
- External oversight, by strengthening accountability (keeping tabs on what agencies are evaluating, how they are conducting evaluations, and their use of evaluations results), can raise the bar on the quality and usefulness of evaluations. An Office of Indigenous Policy Evaluation (established within an independent statutory authority of the Australian Government) could undertake performance monitoring of the Strategy, as well as be a champion for evaluation, innovation and the sharing of evaluation knowledge.
- In line with the Strategy's overarching principle of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges, oversight of the Strategy should have governance arrangements that engage Aboriginal and Torres Strait Islander people. An Indigenous Evaluation Council could provide strategic guidance on the stewardship and monitoring of the Strategy and partner with the Office of Indigenous Policy Evaluation to identify Australian Government-wide evaluation priorities.
- Requiring New Policy Proposals to include evaluation planning, and agencies to publish their rolling Three Year Evaluation Forward Work Plans, could also help embed evaluation into policy planning and policy making processes.

Evaluation approaches and methods are important for conducting quality evaluations, but as Shand argued, ‘the major issues of evaluation are management rather than methodological’ and evaluation policy should be concerned with ‘managing for performance, rather than simply measuring performance’ (1998, p. 15).

Evaluation can only improve the lives of Aboriginal and Torres Strait Islander people (the objective of the Indigenous Evaluation Strategy) if Australian Government agencies use evaluation findings to develop, refine or discontinue policies or programs in the future. As the Organisation for Economic Co-operation and Development (OECD) recently said:

Effective use of evaluations is key to embed them in policy making processes and to generate incentives for the dissemination of evaluation practices. It is a critical source of feedback for generating new policies and developing rationale for government interventions. If evaluations are not used, gaps will remain between what is known to be effective as suggested by evidence and policy, and decision-making in practice. Simply put, evaluations that are not used represent missed opportunities for learning and accountability. (2020a, p. 63)

The Commission consulted with many Australian Government agencies while developing the draft Strategy (appendix A), and while all commented on the importance of evidence-based policy, many senior public servants also said that policy makers do not rely heavily on evidence or past learnings when formulating new policies or refining existing policies. A number of recent reports have also suggested that the Australian Public Service (APS) does not have a strong culture of evaluation and learning.

- The *Independent Review of the APS*, for example, found that Australian Government agencies’ in-house research and evaluation capabilities had undergone a long-term decline, and suggested that the APS needs to rebuild its evaluation capability and practices (DPMC 2019e, pp. 220–221).
- An Australia and New Zealand School of Government (ANZSOG) paper prepared for the Independent Review of the APS concluded that ‘in examining their past performance, departments and agencies are often more concerned with reputational risk, seeking to pre-empt or divert criticism rather than learning from experience and feedback’ (Bray, Gray and ‘t Hart 2019, p. 8).
- The *Independent Review into the operation of the Public Governance, Performance and Accountability Act 2013 and Rule* (‘PGPA Review’) also commented that ‘evaluation practice [had] fallen away’ (though the reason for this was not clear) and needed to be reinvigorated to improve the quality of performance reporting in the APS (Alexander and Thodey 2018, p. 14).

This chapter looks at the governance arrangements needed to support a culture of evaluation and learning within the APS.

- Section 7.1 explores the question: what is a culture of evaluation and learning?
- Section 7.2 looks at ways to enhance the use of evaluations across the APS.
- Section 7.3 looks at what governance arrangements are needed to support high-quality and useful evaluations and encourage experimentation and innovation.

7.1 What is a culture of evaluation and learning?

A culture of evaluation and learning is one that places a high value on evaluation evidence and encourages its generation, dissemination and use (box 7.1). Some of the features of a culture of evaluation and learning are:

- a critical approach to policy and program management, including looking for better ways of doing things, being reflective, and incorporating ‘lessons learned’
- accepting that many answers are not known and being prepared to experiment
- valuing results and evidence
- an organisational environment where it is safe to expose and learn from mistakes and engage in critical self-reflection
- a commitment to transparency, including the publication of evaluations
- leadership that reflects the above qualities.

A culture of evaluation and learning is also one in which the members of an organisation practice ‘evaluative thinking’, as described in chapter 8 (box 8.2).

Box 7.1 A culture of evaluation and learning is one that values evidence

A culture of evaluation and learning is essentially an *evidence-based* culture (there are many examples of public policy literature that explain why evidence-based policymaking is something governments should practice, including several iterations from the Productivity Commission: see, for example, Banks 2009; PC 2010). It requires an agency- and/or government-wide appreciation of the value of undertaking and learning from evaluations, and an institutional framework that prioritises and incorporates evaluation into an agency’s ‘business as usual’ *as part of*, rather than merely *in addition to*, its core policy or service delivery work.

In other words, an agency with a culture of evaluation and learning is one in which staff members understand the purpose and potential benefits of evaluation, and are ‘committed to using findings from evaluation and research to inform decision-making’ (AIFS 2014, p. 3). ‘Learning’ in this context is not only linked to the improvement of individual policies and programs, but is a key plank in the accumulation of organisational knowledge over the longer term.

In that regard, Mayne suggested that an organisation with a strong culture of evaluation:

- engages in self-reflection and self-examination (i.e. self-evaluation) — seeks evidence on what it is achieving; uses the resulting information to both challenge *and* support what it is doing; and values candour, challenge and genuine dialogue
- engages in evidence-based learning — makes time to learn; learns from mistakes and weak performance; and encourages knowledge transfer
- encourages experimentation and change — supports deliberate risk taking and seeks out new ways of doing business (2010, p. 5).

Why is a culture of evaluation and learning important?

A culture of evaluation and learning is a critical factor for an agency to systematically produce high-quality and useful evaluation evidence (the ‘supply side’: Maloney 2017), and to use that evidence when formulating and modifying policies (the ‘demand side’). As the United Nations Evaluation Group said:

Evaluation requires an enabling environment that includes an organizational culture that values evaluation as a basis for accountability, learning and evidence-based decision-making; a firm commitment from organizational leadership to use, publicize and follow up on evaluation outcomes; and recognition of evaluation as a key corporate function for achieving results and public accountability. (2016, p. 13)

An organisation with a strong culture of evaluation and learning will provide its staff with opportunities to improve their evaluation capabilities and develop an evaluative way of thinking (which, in turn, can further strengthen evaluative culture) — chapter 8 discusses capability building in more detail. It will also support high-quality data collection, treat evaluation as an opportunity to improve policies and programs rather than a compliance exercise, and share evaluation results both within and outside the organisation so that they contribute to the greater knowledge commons and can be used by other organisations to benefit future policies, services and research.

A number of participants argued that a culture of evaluation and learning was critical for the success of the Indigenous Evaluation Strategy (the Strategy). Empowered Communities, for example, said:

In terms of adoption of the Indigenous Evaluation Strategy across government, most fundamental to the success of this work-program is buy-in from agencies and centralisation of an evaluation culture in design of policy and programs. Significant effort can be applied to the development of agency evaluation capacity and culture, and tools to enable evaluation processes. (sub. 41, p. 7)

And the Australian Evaluation Society said:

An enabling environment for effective evaluation is a crucial element for the successful implementation and operation of a principles-based Indigenous evaluation framework and identifying evaluation priorities. (sub. 49, p. 27)

Some argued that evaluation needs to become a valued part of agencies’ day-to-day work:

... an [Indigenous Evaluation Strategy] must engender a culture of change within relevant Government departments. This refers to constructive and adaptable attitudes around evaluation and organisational change by mainstream organisations, Government and research bodies ... key stakeholders [should] embrace improvement rather than viewing it as a hindrance to processes and the maintenance of the status quo. (Victorian Aboriginal Community Controlled Health Organisation, sub. 44, p. 10)

Evaluation needs to be integrated into the way we all work – all government departments and agencies need to develop a culture of ‘evaluative thinking’. (National Aboriginal and Torres Strait Islander Health Standing Committee, sub. 104, p. 5)

And the Fred Hollows Foundation pointed out that organisational culture is closely linked to evaluation capacity and capability:

[The Strategy] should include an evaluation capacity building approach, which involves developing an organisational culture of collective learning using ethically and culturally responsive strategies, [thus] ensuring the systems sustainably and continuously support the use of quality evaluation ... (sub. 14, p. 5)

Importantly, the need for this organisational culture is not limited to any particular policy area. With the majority of government spending on Aboriginal and Torres Strait Islander people being for mainstream programs and services (chapter 1), the Strategy will be most effective if it is underpinned by a culture across the APS that values and supports good evaluation and the dissemination of evaluation findings.

7.2 Enhancing the use of evaluations across the APS

While there is strong agreement across Australian Government agencies that policy should be informed by what works, in practice evaluation findings are often not used in policy decision making or to inform governments about where to invest taxpayer funds. An ANZSOG paper prepared for the Independent Review of the APS said:

It is almost universally acknowledged that a systematic approach to high-quality evaluation of policies and programs embedded in APS agencies, as part of a broader approach that reflects on and learns from their own and others' successes and failures, is pivotal to producing evidence to guide appropriate policy settings and effective and efficient public sector management. However, it is equally widely acknowledged that such an approach is not reflected in current APS practice. (Bray, Gray and 't Hart 2019, p. 8)

Results of a recent OECD survey also suggested that one of the main challenges for promoting policy evaluation across government is 'limited use of evaluation results in policy-making', and a conclusion of the report was that 'policy evaluation often constitutes the weakest link in the policy cycle' (OECD 2020a, p. 22).

Many participants to this project expressed a concern that agencies' evaluation activity is often more focused on accountability and compliance than on learning and experimentation — even when, in so many policy areas affecting Aboriginal and Torres Strait Islander people, the solutions are not known. Empowered Communities, for example, said:

We strongly advocate the use of innovative evaluation approaches that allow for adaptive incorporation of evaluation learnings ... there is a need to lift evaluation from being a compliance or academic exercise subject to methodological debate and embed it in a practical and user-friendly way as a learning and adaptation exercise in policy and program delivery. (sub. 41, pp. 9, 12)

The day-to-day demands of government, and the realities of politics (including extensive media coverage), mean that agencies have good reasons to be risk-averse rather than maximising opportunities to learn from innovation, experimentation and public

feedback (see, for example, DOFD 2012, p. 6; Shergold 2015, p. 81) — there can be harsh consequences from getting policies wrong. As the OECD explained:

Innovation runs contrary to the perceived role of bureaucratic organisations. Innovation is new, unknown and risky; by contrast governments have a statutory duty, democratic responsibility and political mandate to deliver public services in consistent and equal ways. Managing these tensions can be complicated for governments, where the risk of innovating appears far greater than the risk of maintaining the status quo. Nor does innovation sit well with the control function of hierarchies which, while they ensure stewardship and accountability over the use of resources, [tend] to discourage risk-taking. (2017, p. 14)

In such an environment it is easy to see why evaluation (and learning from success and failure overall) may take a back seat. The question is, given the realities of policy making, what can be done to embed a culture of evaluation and learning so that government agencies value and use evidence and learning for policy formation?

Conducting useful evaluations

As discussed in chapter 5, engaging with (and taking into account the needs of) evaluation users early in the evaluation process is important for ensuring evaluations have an intended use (or are utilization-focused: box 7.2). Evaluation findings are more likely to be used if they address questions that policy makers want answers for (Haynes et al. 2018, pp. 2–3).

The evidence also suggests that other potential research users — such as service providers — are more likely to seek out and use evaluation findings where evaluators have collaborated and built relationships with those concerned and affected by the policy or program (Patton 2013b). In the context of policies and programs affecting Aboriginal and Torres Strait Islander people, this links to the corresponding need (identified in chapter 5 and also discussed in box 7.5) for the people affected by an initiative to be engaged in the policy design process *before* evaluation, so that their perspectives, priorities and knowledges can be incorporated into the relevant policy design steps. As Patton put it:

[I]ntended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings; they are more likely to understand and feel ownership if they've been actively involved; by actively involving primary intended users, the evaluator is training users in use, preparing the groundwork for use, and reinforcing the intended utility of the evaluation every step along the way. (2013a, p. 90)

The timing of an evaluation also matters. The publication of evaluation findings needs to line up with time periods in which relevant policy decisions are going to be made.

Box 7.2 **Utilization-focused evaluations**

Michael Quinn Patton's concept of utilization-focused evaluations refers to an evaluation planning approach, based on an evaluation being judged on its usefulness to its intended users. Evaluations should be planned and conducted in ways that enhance the likely utilization of both the findings and of the evaluation process itself to inform decisions and improve performance.

Utilization-focused evaluations have two essential elements:

- the primary intended users of the evaluation should be identified and engaged with at the beginning of the evaluation process to ensure that intended uses can be identified
- evaluators should ensure that the intended uses of the evaluation (by the primary intended users as identified above) guide decisions that are made about the evaluation process.

Patton identified a 17-step guide to increase the impact and use of evaluations:

1. Assess and build program and organisational readiness for utilization-focused evaluation.
2. Assess and enhance evaluator readiness and competence to undertake a utilization-focused evaluation.
3. Identify, organise, and engage primary intended users (the personal factor).
4. Conduct situation analysis jointly with primary intended users.
5. Identify and prioritise primary intended uses by determining priority purposes.
6. Consider and build in process uses if and as appropriate.
7. Focus priority evaluation questions.
8. Check that fundamental areas for evaluation inquiry are being adequately addressed: implementation, outcomes and attribution questions.
9. Determine what intervention model or theory of change is being evaluated.
10. Negotiate appropriate methods to generate credible findings that support intended use by intended users.
11. Make sure intended users understand any potential controversies about methods, and their implications.
12. Simulate use of findings (evaluation's equivalent of a dress rehearsal).
13. Gather data with ongoing attention to use.
14. Organize and present the data for interpretation and use by primary intended users: analysis, interpretation, judgment, and recommendations.
15. Prepare an evaluation report to facilitate use and disseminate significant findings to expand influence.
16. Follow up with primary intended users to facilitate and enhance use.
17. Conduct meta-evaluation of use: be accountable, learn and improve.

Sources: BetterEvaluation (2011); Patton (2013b).

Making evaluation evidence easy to access

For policy makers, researchers, Aboriginal and Torres Strait Islander organisations and others to make use of evaluation evidence, they need to know about it. As the OECD said:

The first step to promote use is therefore that the results be made available to their intended users — simply put, that they be communicated and disseminated to stakeholders. (2020a, p. 105)

The Commission's review of the evaluation evidence base for policies and programs affecting Aboriginal and Torres Strait Islander people found that there is evaluation evidence already available that policy makers, and other stakeholders, may not be aware of. Even if stakeholders *are* aware of such evidence, it is often not easy to find or to collate.

There can be a high price for policy makers not being aware of, or not having easy access to, evaluation and research findings. In the United Kingdom, the Nesta Foundation's Alliance for Useful Evidence argued that had policy makers put more weight on the clinical research around cot death, many babies' lives could have been saved — a 2005 meta-analysis reported that by 1970 there was significant evidence that putting babies to sleep on their fronts increased the risk of cot death compared with putting babies on their backs, but the safer sleeping position was not consistently recommended in public health campaigns until 1995 (Alliance for Useful Evidence 2016, p. 12; Gilbert et al. 2005).

Publishing evaluation reports is key to making evaluation evidence easy for program managers and policy makers (and other evaluation users) to access and use. Many participants commented on the importance of publishing evaluations to promote the use of evaluation findings (box 7.3).

While some Australian Government agencies — including the National Indigenous Australians Agency (NIAA) — have a policy of publishing evaluations, the majority do not, and many evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people undertaken by Australian Government agencies have not been made public (chapters 2 and 3). The Victorian Aboriginal Child Care Agency pointed out the problems with this varied approach:

Practice in relation to evaluation transparency is mixed, with some government agency external evaluation reports published and/or made available to implementing agencies, but in some cases findings not made available. This undermines a key aim of evaluation — to enable those designing and implementing services to continuously improve service delivery. (sub. 26, p. 8)

The Independent Review of the APS also noted the importance of publishing evaluations (and other research) to maximise transparency and share the benefits of new evidence. It recommended that agencies should publish evaluations by default, unless exempted from this requirement by Cabinet (DPMC 2019e, p. 223). In response, the Australian Government stated that it supported the 'appropriate publication of completed evaluations' (DPMC 2019d, p. 22), but did not specify whether or not this would require publication.

Box 7.3 **Strong support for making evaluation findings available**

Empowered Communities:

... supports the publishing and public distribution of evaluation reports except when sensitive cultural information is included or if data privacy and confidentiality conditions cannot be assured. (sub. 41, p. 12)

The National Mental Health Commission:

... it will be essential to ensure that the findings of evaluation processes are made widely available. The Commission considers that the Strategy should incorporate a presumption in favour of making the findings of evaluation processes publicly available. (sub. 19, p. 3)

Tony Dreise et al.:

Given the focus the Government is placing on evaluation as a driver of better outcomes, there is an overwhelming argument in favour of mandating the public release of all evaluations within two months of their finalisation. (sub. 33, p. 6)

Victorian Aboriginal Community Controlled Health Organisation:

The report that is produced must be published and shared with all evaluation participants and relevant Communities that fall under similar policies or programs that were evaluated. ... Reporting must be written in plain English to ensure it is accessible to all Community members. (sub. 44, p. 21)

Indigenous Allied Health Association:

Governments and commissioning agencies should be required to publicly disseminate evaluation findings, including poor performance, wherever practicable. This is essential for accountability and transparency, including in the use of public resources, as well as to ensure we can learn from what did not work. Too often, even where evaluations are undertaken, we fail to fully implement the recommendations or act on areas for potential improvement in service delivery. (sub. 31, p. 4)

Australian Institute of Aboriginal and Torres Strait Islander Studies:

Publicly available evaluation reports are important for transparency and accountability of government spending, but also to build the knowledge and evidence base upon which Indigenous peoples and organisations can make decisions. ... There may be information contained in the report that is culturally sensitive, commercially sensitive or have legal ramifications. In such circumstances it may be appropriate to redact or publish a summary of the report. This means that consent protocols in the evaluation framework need to include potential publication of materials. (sub. 72, p. 17)

National Indigenous Australians Agency:

Sharing evaluations helps inform design, delivery and decision-making. Under the [Indigenous Advancement Strategy] Evaluation Framework, all evaluation reports or summaries are to be made publicly available. In cases where ethical confidentiality concerns or commercial in confidence requirements trigger a restricted release, summaries of the findings will be published in lieu of a full report. (sub. 86, p. 3)

The Western Australian Government:

Even excellent evaluations are of limited benefit if their results are not made available to decision-makers and stakeholders so that future decisions can be informed by evidence. ... Where possible, evaluation reports should be published for reasons of accountability and transparency, but this will not be necessary or appropriate in all instances. (sub. 74, pp. 7, 9)

Indigenous Community Volunteers:

Critical to the publication of results is creating a 'fail forward' environment across agencies to learn, accepting that all programs will have failures and room for improvement. ... To overcome the lack of transparency in programs and evaluations, the Strategy would need to mandate an expectation regarding publication and formal management responses to evaluations. (sub. 88, p. 20)

As well as increasing the use of evaluation results in policy formulation and modification, requiring evaluation results to be made public can:

- provide an incentive to improve the quality of evaluations (because evaluators know that the report will be made public and that others can critique the evaluation approaches, methodologies, and analyses used)
- allow more informed public debate about the effectiveness (or otherwise) of policies and programs
- increase agencies' accountability to Aboriginal and Torres Strait Islander people, and to the public overall (while accountability needs to be balanced with the pursuit of learnings, it is also an important goal in its own right: Cabaj 2017)
- increase the pressure for more systematic follow-up of evaluation findings.

Further, the Lowitja Institute has previously argued that 'transparency in evaluation is an ethical requirement, as Aboriginal and Torres Strait Islander communities have the right to know about the effects of programs that affect them, and a right to the findings from evaluations they have participated in' (Kelaher et al. 2018b, p. 25).

Publication of all (or most) evaluation findings is relatively common among other OECD countries. Just under half of all OECD member countries make evaluation findings and recommendations publicly available by default (OECD 2020a, p. 106). In Canada, for example, evaluation reports are published on agencies' websites, while in Norway, evaluations carried out on behalf of government agencies are published in a single publicly-accessible online evaluation repository (appendix D). As discussed in chapter 3, the results of our information request suggest that only about 10 per cent of responding agencies always made evaluations public.

Could requiring evaluations to be published affect agencies' behaviour?

Requiring evaluation reports to be made public could, however, undermine the incentive for Australian Government agencies to undertake evaluations. Some participants suggested that it could also influence the content that agencies are prepared to include in evaluation reports (and there was a risk that reports would be progressively sanitised as they moved through an agency's management hierarchy for publication approval).

A number of departments told the Commission about a reluctance to publish evaluations when policies or programs had been found to be ineffective or not working as intended. Liz Curran also reported requests for 'independent evaluators to change key findings' and for words 'to reflect better on the Ministry' (sub. 2, p. 3).

The reality is that no agency will find it easy to disclose negative or controversial evaluation findings. But when policies are found to not work — or it becomes apparent that the costs of a policy outweigh the benefits, or that the policy is causing harm — it is important that

this information is shared: the greatest tragedy of failure is failing to learn from it (Banks 2013, p. 18).

An agency's credibility could also be enhanced by demonstrating a willingness to publish evaluations with both negative and positive findings (especially if the agency's management actively acknowledges and responds to those findings, explains what they have learned from the evaluation process, and presents a management plan for how the findings will be incorporated into the agency's work going forward). The Independent Review of the APS concluded that:

One challenge to reversing the decline in evaluation work across the APS is the risk of failures (or at least failure to meet all expectations) being exposed. The counterpart, as one former secretary put it, is that 'successes can also be celebrated and hopefully people's trust in government will also be enhanced by greater honesty in making available objective and rigorous assessments of performance'. (DPMC 2019e, p. 222)

Trust between governments and Aboriginal and Torres Strait Islander people, communities and organisations is particularly important in light of Australia's colonial history and its continuing impacts (chapter 1).

As a number of participants pointed out (box 7.3), however, there may also be circumstances where it is objectively not appropriate or desirable to publish evaluation reports — for example, where doing so would compromise confidentiality or privacy, or where there is culturally sensitive information. In such cases, a summary report could be made public instead (this is the approach taken by the NIAA: DPMC 2018b, p. 12). The summary report would need to include an explanation about why the full report could not be made available.

Weighing up the benefits of making evaluations publicly available against the risks and costs, it is the Commission's view that there should be public reporting requirements for evaluations under the Indigenous Evaluation Strategy (draft Strategy, action 7). Where there are compelling concerns that publishing the full evaluation report would compromise confidentiality or privacy, or where there is culturally sensitive information, a summary report should be published instead (draft Strategy, action 7). If these concerns only apply to a part of an evaluation, the remainder of the evaluation should be made public.

Evaluations or summaries should be published no more than three months after completion, and they should also be made available to a central evaluation clearinghouse (draft Strategy, action 9; discussed in more detail in the next section).

Building a cumulative body of evaluation evidence

Good knowledge management is also critical for the systematic use of evaluation findings across Australian Government agencies, and it is missing under current arrangements

(chapter 2). One of the findings of a paper prepared for the Independent Review of the APS was that:

At the central level, even when evaluations are undertaken, there is no standard mechanism for the findings to be considered and addressed. Nor are there strong mechanisms at the central or departmental level for ... bringing the findings together in a strategic way. In addition, there are no strong mechanisms to build a history of what has worked and what has not, or to systematically review evaluations to learn cross-cutting lessons and provide guidance to future evaluation efforts. (Bray, Gray and 't Hart 2019, p. 15)

Evidence from evaluation is a public good, in that when it is disseminated anyone can benefit from it without taking away from the benefits to others. Developing a cumulative body of knowledge from evaluation, upon which users across the public, private and non-profit sectors can draw, is therefore critical for maximising the benefits of evaluation.

A number of other countries have national databases or evaluation portals to centralise evaluation evidence in one easily-accessible place. The World Bank, commenting on the importance of good processes for disseminating evaluation results, said:

Faster and easier access to relevant knowledge from evaluations can help staff work more efficiently and effectively. Establishing simple and systematic processes for knowledge management to leverage the vast body of knowledge from evaluation — and effectively capture, curate, share and apply such knowledge — is critical to improving results on the ground. (World Bank 2019, p. 3)

As is the case in most areas of social policy, no single evaluation will provide robust information that can be generalised, but the information from an evaluation *adds* to the evidence base built up over time and across different contexts. The reliability of certain types of policy thus increases when there are a reasonable number of consistent evaluations (OMB 2019, p. 59). As Weiss put it:

As hundreds and thousands of evaluations are done on similar kinds of programs, evidence mounts up ... Sometimes the evidence is synthesized through meta-analysis, which gives an overall estimate of the outcomes of programs of similar kind; sometimes the synthesis comes through qualitative reviews. ... When evaluation adds to the accumulation of knowledge, it can contribute to large-scale shifts in thinking — and sometimes, ultimately, to shifts in action. (1998, p. 24)

The importance of building the evidence base by developing a central repository for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people was noted by many participants to this project (box 7.4). Some referred to the Closing the Gap Clearinghouse — a Council of Australian Governments initiative to collect, analyse and synthesise evaluation and research evidence on what worked to close the gap in Indigenous disadvantage (which was in place from 2009 to 2014: chapter 3) — and the need for a similar central repository of evaluation evidence. Tony Dreise et al., for example, said:

... the Closing the Gap Clearinghouse is no longer functioning. Its closure has created a gap in the Indigenous Evaluation and related research space. (sub. 33, p. 6)

There is also a growing body of evaluation evidence generated by Aboriginal and Torres Strait Islander organisations, non-government organisations (NGOs) and state and territory governments that could be better drawn on if it was easier to find.

Box 7.4 Calls for a central repository and strategies for sharing information

The Fred Hollows Foundation:

The Indigenous Evaluation Strategy should ensure evaluation results are used to inform and guide policy and program design by: ... Developing a public database of all evaluations commissioned by the Australian Government in relation to policies and programs impacting Aboriginal and Torres Strait Islander Peoples and report on how the learnings from the evaluation have been used. (sub. 14, p. 2)

First Nations Media:

We encourage the Productivity Commission to seek efficiencies in data collection and evaluation through a centralized repository for information that is available to the public and has a closed area for inter-departmental information sharing. (sub. 30, p. 10)

Save the Children:

The evidence should be available in a single place, which includes existing resources such as those prepared for the Closing the Gap Clearinghouse and new research. It should include underlying data in a form that can be analysed and disaggregated at least to program or community level. ... It should also include brief summaries and evidence reviews, as well as more detailed year-by-year data reports. (sub. 52, p. 3)

Generation One, Minderoo Foundation:

Effective strategies must be developed to share evaluation information. (sub. 5, p. 5)

Tony Dreise et al.:

There is a clear need for an agency — perhaps the Productivity Commission — to centrally archive and make available all evaluations and reviews of policies and programs affecting First Nations people. (sub. 33, p. 6)

Victorian Aboriginal Child Care Agency:

Evaluation findings should be supporting the development of an overall evidence-base of Aboriginal programs and policies. This will not happen if evaluation findings are not published or made available to Aboriginal organisations. (sub. 26, p. 8)

Some organisations already operate (or are developing) repositories, clearinghouses or ‘hubs’ for research relating to Aboriginal and Torres Strait Islander people.

- The Australian Indigenous HealthInfoNet provides support for people working in the Aboriginal and Torres Strait Islander health sector by making research and other knowledge readily accessible.
- All research that takes place under the auspices of the Lowitja Institute (which relates to Aboriginal and Torres Strait Islander health and wellbeing, and is primarily led or coordinated by Aboriginal and Torres Strait Islander community controlled organisations) is published in full on the Institute’s website.
- The Australian Institute of Health and Welfare (AIHW) 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: sub. 99, p. 5) and an Indigenous Mental Health and Suicide Prevention Clearinghouse. The AIHW said the newly established Indigenous Mental Health and Suicide Prevention Clearinghouse, together with the Indigenous Evaluation Strategy, 'presents opportunities for collaboration in this area' (sub. 99, p. 1).

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) is also in the early stages of developing an Indigenous Research Exchange, which aims to provide a portal for research by Aboriginal and Torres Strait Islander academics, translate research into 'useful and practical information', identify gaps in the evidence base, and provide grants for research priorities to fill those gaps (AIATSIS 2020; Payne, Chi-Chung and Knight 2019).

'Making sense' of evaluation evidence through translation and synthesis

However, publication alone is unlikely to significantly improve the uptake of evaluation in policy making. The OECD suggested that national databases on their own could even hinder the use of evaluations:

Portals serving as passive repositories of information are less likely to promote evidence use ... Compiling evaluations in portals or databases runs the risk of information overload, thus hindering the incorporation of findings and reducing the effectiveness of evaluation. As the number of evaluations increases, it becomes more difficult for policy makers and practitioners to keep abreast of the literature. ... In addition, such repositories do not necessarily allow stakeholders to understand the quality of the evidence produced by an evaluation: its rigor or replicability for example. (OECD 2020a, p. 108)

Rather, it is when evaluation results are well-synthesised, tailored for specific users and disseminated to users appropriately (including the communities and organisations that participate in evaluations: Indigenous Community Volunteers, sub. 88, pp. 20–21) that uptake of evaluation in policy making is more likely (Maloney 2017, pp. 33–34).

Some countries seek to enhance the usability and usefulness of evaluation findings by contextualising and/or synthesising the findings. Evidence synthesis methodologies aggregate evaluation findings and review them in a systematic way (and generally also assess and rate the strength of the evidence). Evidence synthesis is a useful dissemination tool as it helps policy makers access large bodies of evidence while also having a sense of the quality of the evidence.

The United Kingdom's *What Works Network* is one example of a government initiative to improve the sharing and application of research evidence. The Network currently has nine Centres (each with a focus on a unique policy area), but all Centres share a common remit that includes knowledge translation (appendix D). The Centres produce policy briefs to disseminate key messages to their target audiences. There are also some large international non-profit organisations and networks pursuing a similar goal of synthesising research to increase its use by decision-makers, including Cochrane (which focuses on health research)

and the Campbell Collaboration (a sister initiative of Cochrane, with a broader remit of economic and social policy research) (Campbell Collaboration 2020; Cochrane 2020).

Collecting, disseminating and synthesising evaluation findings — collectively referred to as ‘knowledge translation’ — will be critical for improving the use of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. Any one policy or program evaluation could be relevant to the work of several Australian Government agencies, Aboriginal and Torres Strait Islander organisations, or service providers, and could relate to several different fields of existing evidence. Leaving knowledge translation entirely to individual evaluation teams or agencies therefore risks a siloed approach to evidence generation and use.

Having a central knowledge translation function for evaluations under the Strategy would help to mitigate this risk, and optimise the sharing of evaluation evidence. That said, evaluators and agencies will need to *participate* in the knowledge translation process, with a core requirement of this being to engage with the Aboriginal and Torres Strait Islander people affected by a policy or program (box 7.5).

However, for the most effective sharing of knowledge, these functions would be set up to cover all Australian Government social and health policy evaluations, not just evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. The Commission is proposing that, over the longer term, a new independent Australian Government agency — a Centre for Evaluation Excellence (section 7.3) — be given responsibility for collecting, disseminating and synthesising the findings of Australian Government social and health policy and program evaluations.

In the interim, a Clearinghouse for evaluations conducted under the Indigenous Evaluation Strategy should be established within an existing independent statutory authority. There are a number of agencies that could take on this role:

- The Australian Institute of Health and Welfare and the Australian Institute of Family Studies (AIFS) both have experience in operating the Closing the Gap Clearinghouse, and synthesising research related to Aboriginal and Torres Strait Islander people. The AIHW currently has the largest public repository of research on improving outcomes for Aboriginal and Torres Strait Islander people. AIFS has also demonstrated skills in making sense of evidence and on knowledge translation (through, for example, its *Child Family Community Australia* information exchange).
- The Australian Institute of Aboriginal and Torres Strait Islander Studies has an exclusive focus on research involving Aboriginal and Torres Strait Islander people (and strong relationships with Aboriginal and Torres Strait Islander organisations) and is responsible for the Guidelines for Ethical Research in Australian Indigenous Studies (GERAIS). AIATSIS has also recently launched its own Indigenous Research Exchange.
- The Productivity Commission, as suggested by some participants, could also take on the role. Because of our inquiry and research work, we have experience in synthesising research from many disparate sources, and our ongoing reporting work (such as the

Overcoming Indigenous Disadvantage series) means we have experience working with Aboriginal and Torres Strait Islander people and organisations.

Box 7.5 What does knowledge translation entail in practice?

Knowledge translation (or research translation) is ‘an interactive interchange of knowledge between research users and researcher producers ... [Its primary purposes] are to increase the likelihood that research evidence will be used in policy and practice decisions and to enable researchers to identify practice and policy-relevant research questions’ (Mitton et al. 2007, p. 729).

The Lowitja Institute argued that knowledge translation is not simply a process of disseminating the final results of research, but rather a dynamic series of interactions that affect every step of the research process, allowing research practitioners to create direct links between community priorities, the evidence base, policy development, implementation, and monitoring/evaluation (Smith 2018). Bourassa (2017, as cited in Smith (2018)) also argued that ‘if you don’t have community engagement [and] Indigenous people in the driver’s seat, you don’t have Indigenous knowledge translation’.

The following table summarises some knowledge translation strategies that the Lowitja Institute implements in its work, and provides examples of how these might be applied in the context of the Indigenous Evaluation Strategy.

Knowledge translation strategies at each step of the research/policy process

Stage	Strategy	Examples
Policy/program design	Consider Aboriginal and Torres Strait Islander peoples’ policy/research priorities	Agencies develop new initiatives/pilots with Aboriginal and Torres Strait Islander people, communities and organisations (see, for example, Winnunga Nimmityjah AHS, sub. 46, p. 2).
Evaluation planning	Build community- and policy-relevant questions	Include Aboriginal and Torres Strait Islander people/communities/organisations in early design of evaluation processes, indicators, and mechanisms for sharing and applying learnings (Empowered Communities, sub. 41, p. 11).
Conducting evaluation	Strengthen community capacity to understand and apply results	Share results with participating communities and organisations <i>before</i> report publication, enabling participants to contribute to the interpretation of results (AIMN, sub. 15, p. 13).
Publishing/sharing results	Interpret results for policy relevance; communicate appropriately to stakeholders	Disseminate results through interactive knowledge exchange with: stakeholder forums (cross-agency); research and policy networks/communities of practice; Aboriginal and Torres Strait Islander community controlled organisations.
Post-publication	Integrate results into evidence base	Highlight results in relevant journal articles, systematic reviews, policy briefs, fact sheets, media campaigns; facilitate policy dialogues, coverage at conferences.
Policy/program refinement and continuing practice	Implement results-driven change; redefine research/ evaluation questions	Agencies produce action plans for necessary changes in programs/policies/systems and are held accountable to these plans; monitoring and evaluation approaches are refined to reflect these changes.

Sources: Adapted from Edwards, Zweigenthal and Olivier (2019); Lowitja Institute (2020); Smith (2018).

The Commission would like to hear participants' views about which existing Australian Government agency would be best placed to collect, disseminate and translate evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

INFORMATION REQUEST 7.1

The Commission is seeking participants' views on which current Australian Government agency would be best placed to house the Indigenous Evaluation Clearinghouse: the Australian Institute of Health and Welfare; the Australian Institute of Family Studies; the Australian Institute of Aboriginal and Torres Strait Islander Studies; the Productivity Commission; or some other agency?

Format also matters for enhanced use

Demand for evaluation findings can also be increased by making them available in forms that are understandable and usable for a wide variety of audiences. In Canada, for example, the *Standard on Evaluation* (a component of the *Directive on Results*) stipulates that all evaluation reports must meet certain requirements around readability, including: providing an accurate assessment of the contribution of the program to its related government priorities and/or departmental results and priorities; presenting appropriate context for the evaluation and the policy, program, priority, unit or theme being evaluated, including by identifying the limitations of the evaluation in a way that informs readers about the reliability of findings and conclusions; and making clear, actionable recommendations that aim to address the key issues or concerns identified (Government of Canada 2016a).

While all evaluation reports should be written in plain English, a concise, simple language summary would be useful for sharing evaluation results more widely, recognising not only that those results are relevant to a wide range of audiences but also that it is not always possible to avoid technical language in a report itself.

The addition of a brief, clear summary to all evaluation reports is in line with the National Health and Medical Research Council's (NHMRC) suggestion in *Keeping research on track II* — it recommends that all primary research with Aboriginal and Torres Strait Islander communities should produce a Plain Language Community Report for sharing results with research participants (NHMRC 2018b, p. 19).

Several participants highlighted the importance of reciprocity in Aboriginal and Torres Strait Islander relationships (chapter 1) and of delivering research results back to Aboriginal and Torres Strait Islander communities, including in community languages where appropriate, to provide accountability and share the benefits of research.³⁰ Sharing research results with participants is also an important aspect of ethical research and accountability more broadly

³⁰ See, for example, AHCWA (sub. 42, p. 3); AIMN (sub. 15, p. 13); Central Australian Aboriginal Congress (sub. 48, p. 6); Sophia Couzos (sub. 92, p. 4); First Nations Media Australia (sub. 30, p. 10).

(see, for example, the *National Statement on Ethical Conduct in Human Research*: NHMRC, ARC and UA 2018, pp. 38–40).

And as the Western Australian Government said, publication of reports is not always the most appropriate method of communicating evaluation findings (including to Aboriginal and Torres Strait Islander communities); other appropriate methods might include community workshops, information sessions or verbal feedback (including in community languages where necessary: sub. 74, p. 10). Engaging with Aboriginal and Torres Strait Islander people and organisations early in the evaluation about how results are best shared is therefore important (chapter 5).

But it is not just evaluation results that can provide insights for evaluation users — there can also be important lessons about evaluation commissioning, approaches and methodologies. For example, the former Department of Infrastructure and Regional Development (now the Department of Infrastructure, Transport, Regional Development and Communications), in its Evaluation Strategy 2016–2021, requires evaluation reports to be followed up with a ‘Lessons Learned’ document (DIRD 2016, p. 10) that debriefs Department staff on key learnings from an evaluation, including comments about the strengths and challenges of the evaluation process itself. Disseminating knowledge about good evaluation practices within (and potentially across) agencies can also be important for capability building (chapter 8).

With this in mind, together with the need to monitor agencies’ performance against the Strategy and the impact of the Strategy on policy and program design more generally (chapter 10), as well as the importance of transparency of evaluation processes for accountability, the Commission is proposing that all evaluation reports under the Strategy should document how an evaluation adhered to the Strategy’s principles (draft Strategy, action 8). Further, agencies’ management responses (which could be included as part of the evaluation report or published separately, but in any case should be published in a timely fashion following the completion of the evaluation) should include an explanation of what they have learned from the evaluation process, what they have changed in response to the findings, and any further action they intend to take.

Summarising what is needed to maximise the accessibility and useability of evaluation findings

To increase the likelihood of policy makers incorporating evaluation findings into future policy and program design, the results of evaluations carried out under the Indigenous Evaluation Strategy (the Strategy) need to be easier to access, understand in context, and use. (This would also have the benefit of better enabling researchers, service providers, Aboriginal and Torres Strait Islander organisations and other stakeholders to incorporate evaluation results into their own work.)

There are **three** actions that would feed into better accessibility and useability:

1. All reports of evaluations under the Strategy should be published and made available on agencies' websites within three months of being completed, unless there is a compelling reason to do otherwise (such as where publishing the report would compromise confidentiality or privacy, or where there is culturally sensitive information). If this is the case, a summary of the evaluation's findings should be published, including an explanation for why the full evaluation report could not be published (draft Strategy, action 7).
 - (a) The responsible Australian Government agency should publish a management response to each evaluation or summary report (draft Strategy, action 10). This response should explain what they have learned from the evaluation and what actions they have taken, or propose to take, in response to the findings.
 - (b) Published evaluations should also be made available to the Indigenous Evaluation Clearinghouse (see below) for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.
2. The Australian Government should establish an online evaluation evidence clearinghouse — the Indigenous Evaluation Clearinghouse — for evaluations of policies and program affecting Aboriginal and Torres Strait Islander people (draft Strategy, action 9). Initially, the Clearinghouse should sit within an existing independent statutory authority; as the Commission has not formed a clear view on which agency this should be, participants' views on this question are welcomed.
 - (a) The statutory authority responsible for the Clearinghouse should also be responsible for coordinating the synthesis of new evaluation evidence into the existing evidence base, and for translating the knowledge into forms that are accessible to a variety of different audiences. If the Centre for Evaluation Excellence is established (section 7.3), the Indigenous Evaluation Clearinghouse should then operate as a section within the whole-of-Australian Government policy and program evaluation repository.
3. All evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people undertaken or commissioned by Australian Government agencies should include a short plain English summary of the evaluation report. Other ways of sharing evaluation findings, such as community language translations, verbal feedback, or information sessions, should be agreed on with the relevant Aboriginal and Torres Strait Islander people, organisations and communities at the evaluation planning stage. Reports of evaluations under the Strategy (or summaries of findings) should also document how the planning, commissioning and conduct of the evaluation adhered to the principles of the Strategy (draft Strategy, action 8).

7.3 Creating incentives to produce high-quality and useful evaluations

Australian Government agencies will only produce high-quality and useful evaluations if they have an incentive to do so. As Maggie Walter said:

Behavioural change only comes when people have skin in the game through some measure of accountability or responsibility for the outcomes of their actions. There is no shortage of goodwill in departments but that is not enough. The public service needs a reason to move beyond what it is doing now, to consider and include evaluation in a more comprehensive sense, which will only come with greater incentivisation. (sub. 112, p. 1)

As discussed in chapter 2, Australian Government agencies are required under the Enhanced Commonwealth Performance Framework to report, through Annual Performance Statements, on results achieved against the measures set out in Corporate Plans and Portfolio Budget Statements. A number of agencies told the Commission that the Framework, with its new requirements for strategic planning, measuring and assessing performance, and reporting, had meant greater external scrutiny. However, there are no formal requirements for evaluation under the Framework (chapter 2), and some participants argued that accountability is currently focused more on demonstrating that funds have been used appropriately than on the effectiveness of government spending on policies and programs.

The recent ANZSOG paper, prepared for the Independent Review of the APS, explained that the reality of policy making is that:

In their day-to-day activities, departments and agencies tend to focus on achieving the ‘here and now’ priorities of program and policy implementation and of ministers. ... In this context, evaluation is often also seen as yesterday’s news or a second-order and lower-priority issue. Evaluation processes and findings then become just one more thing that needs to be defensively managed, or an opportunity for quick grabs to justify current and future decisions and activities. (Bray, Gray and ‘t Hart 2019, p. 8)

And under current arrangements, the only external scrutiny of Australian Government agencies’ evaluation processes is that undertaken by the Australian National Audit Office (ANAO, which recently released a report on performance reporting and evaluation under the Indigenous Advancement Strategy Framework)³¹ and, at times, by the Commission (we recently made recommendations on evaluation in *A Better Way to Support Veterans* (PC 2019a), *Mental Health (Draft Report)* (PC 2019b) and *Expenditure on Children in the Northern Territory* (PC 2020)).

³¹ The ANAO also concluded, in an audit of the Mandatory Minimum Requirements (MMRs) for Aboriginal and Torres Strait Islander businesses’ participation in procurement under the Indigenous Procurement Policy, that a lack of centralised monitoring and oversight had undermined the MMRs’ implementation (ANAO 2020, p. 8).

External oversight to change incentives

One way to strengthen accountability is to have external oversight. By keeping tabs on what agencies are evaluating, how they are conducting evaluations, and their use of evaluation findings in policy and program development, this can encourage agencies to conduct evaluations that are useful and of high quality, and to make greater use of those evaluations (chapter 2; appendix D). As the OECD said:

While individual [evaluation] competencies are important, formal organisations and institutional mechanisms set up a foundation for evidence-informed policy making that can withstand transitions between leadership ... The use of evaluations is intimately linked to organisational structures and systems, insofar as they create a fertile ground for supply and demand of evaluations to meet. (2020a, p. 116)

The Letter of Direction for this project points to the Productivity Commission playing an oversight role for the Strategy. It notes that the Commission will ‘review the performance of agencies against the Strategy over time, focusing on potential improvements and on lessons that may have broader application for all governments’. The Treasurer, on announcing the Commission’s Indigenous Policy Evaluation Commissioner, also spoke about ‘the Commission’s enhanced role in Indigenous policy and program evaluation’ (Frydenberg 2018).

A number of participants argued that an independent oversight body was required for the Strategy to be effective. Some participants supported the Productivity Commission taking on this role, although some proposed that the Commission be augmented by Indigenous representation or an Indigenous Council. Others argued for the oversight body to be Indigenous-led, or for the oversight role to be allocated to a non-government Aboriginal and Torres Strait Islander body (box 7.6). Other suggestions were that the oversight role sit with AIATSIS, the ANAO, or the Australian Human Rights Commission (AHRC).

Each of the proposed statutory bodies have the independence from the day-to-day activity of government necessary to objectively assess agencies’ performance against the Strategy (chapter 10). While the Productivity Commission’s experience in performance monitoring means it is well-suited to undertaking a performance assessment and monitoring role, the ANAO is also well placed to perform such a role. The Commission’s consultative and transparent approach to inquiries and research would also be important for the monitoring role, but other agencies also operate in a similar way (the AHRC, for example, undertakes extensive consultation in carrying out its remit).

It is important to note that none of the abovementioned agencies would have powers to directly influence agencies in the exercise of their day-to-day functions (rather, influence would be created by the advice given to agencies and by reporting on their performance against the Strategy). It would be transparency of agencies’ performance that would provide an incentive for agencies to lift the bar on the quality and usefulness of their evaluations.

Box 7.6 External oversight for the Strategy: what participants said

Tony Dreise et al.:

... an independent First Nations governance structure should be given oversight of Indigenous evaluation. Ideally this new structure would be resourced to commission or undertake evaluations directly. Either way, a body with the power to oversee evaluation is needed (rather than merely being an advisory or consultative body). First Nations control of evaluation would bolster the confidence of First Nations peoples in the Commonwealth's administration of Indigenous Affairs. ... ensuring substantive compliance with [the Strategy] will present a considerable challenge without an independent evaluation body. ... There are ... strong arguments in support of exploring a new approach that takes at least some evaluation selection, planning and implementation outside of agencies. (sub. 33, pp. 3, 6)

The Department of Health:

Health envisages that the Commission could take a leadership role on evaluation activities in the future. This role could be used to identify and evaluate key areas of cross-cutting work across specific policies and programs affecting the health, social and economic outcomes of Aboriginal and Torres Strait Islander people and communities. The Commission's role could encompass evaluations undertaken across government, with priorities set in consultation with other agencies, and with Aboriginal and Torres Strait Islander people and communities. (sub. 35, p. 4)

Michael Dillon:

[M]y intuition tells me that a part of the solution would be the establishment of a standalone and statutorily independent Indigenous evaluation entity. ... The appointment of a Commissioner or Commissioners should be done with Indigenous input and to the extent possible made non-partisan to ensure independence. (sub. 16, p. 5)

Aboriginal Medical Services Alliance NT (AMSANT):

AMSANT is strongly in support of the establishment of a national Aboriginal and Torres Strait Islander body with direct input into the Federal Parliament, in line with the Uluru Statement from the Heart. We envisage that such a body would have a key role in both the development and review of policies and projects relating to Aboriginal people. ... However until the establishment of a 'voice' to parliament, we support an ongoing role for the Productivity Commission in overseeing the implementation of an Evaluation Strategy. It is important that this mechanism remains independent from Government and could be achieved by establishing an Indigenous Evaluation Council within the Productivity Commission. (sub. 81, p. 7)

Local Government Association of Queensland:

The Australian Government should assign an independent body to oversee and report on the operation of [the Strategy] for remote and discrete communities. ... These functions may be allocated to an existing organisation, but should be established in legislation and include appropriate expertise and Indigenous representation. (sub. 102, p. 12)

Lowitja Institute:

The Strategy should include the establishment and resourcing of one non-government agency to coordinate and facilitate evaluation activities with government departments to:

- support and develop evaluation and quality improvement literacy;
- ensure a consistent approach to evaluation across departments;
- monitor and evaluate practice; and
- translate evaluation outcomes more broadly into policy and practice. (sub. 50, p. 10)

The Western Australian Government:

One option could be for a national evaluation support body (which might sit within the Productivity Commission) to be tasked with providing direction, support and leadership to agencies in monitoring and ongoing improvement in evaluation, including training in evaluation standards. This body could also conduct reviews of agency evaluations [against] the Strategy. (sub. 74, p. 11)

Promoting better practice, not just compliance

Strong leadership is important for building an organisational culture that values evidence and innovation, builds evaluative capability, shares knowledge (including about what has not worked) and learns from the results of evaluations (see, for example, Mayne 2010, pp. 8-12). As Alexander and Thodey (2018, p. 13) said, ‘the tone is set at the top’. In a survey by Maloney (2017, p. 30), most respondents also indicated that leadership had more of an impact on evaluation use than agency evaluation policies and processes.

For the Strategy to be most effective, a central evaluation function should not only monitor compliance — it should also provide guidance to agencies on implementing the Strategy, and be a ‘champion’ for the value of evaluation (including by helping agencies to view evaluation less as a compliance exercise, and more as a key tool for finding innovative solutions to complex policy problems³²).

Ideally, the central evaluation functions would all be located in one place. This points to the need for a standalone unit within an existing independent Australian Government statutory authority — we have called it an ‘Office of Indigenous Policy Evaluation’ (OIPE) — to house these functions. The proposed OIPE would:

- have stewardship of the Strategy
- champion the value of evaluation in policy making and identify potential areas for evaluations or systematic reviews on cross-cutting topics
- provide advice to Australian Government agencies on what the Strategy means in practice
- identify evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people (chapter 6)
- oversee agencies’ rolling Three Year Evaluation Forward Work Plans (chapter 6)
- coordinate with the Head of Evaluation Profession on capability building opportunities (chapter 8)
- monitor and publicly report on Australian Government agencies’ performance against the Strategy, and on how effective the Strategy has been in improving the quality and usefulness of evaluations affecting Aboriginal and Torres Strait Islander people (chapter 10).

The OIPE would not be expected to undertake policy or program evaluations. As discussed in chapter 5, evaluation should be integrated into the policy making process, which means agencies are best placed to undertake and commission such evaluations. That said, the OIPE could identify and make recommendations to the Australian Government on, and undertake,

³² As discussed in chapter 3, some participants raised concerns about evaluations being more focused on accountability (demonstrating value for money and that programs have achieved their stated objectives via performance reporting) and less on finding ways to improve outcomes for Aboriginal and Torres Strait Islander people.

evaluations or reviews of outcome areas that cut across more than one agency or policy area. These would involve collaborating closely with policy and program agencies and key stakeholders to work through complex questions of various policy and program impacts on Aboriginal and Torres Strait Islander people.

The central Indigenous Evaluation Clearinghouse and knowledge translation function could also be located in the same agency as the monitoring, advisory and priority-setting functions.

INFORMATION REQUEST 7.2

The Commission is seeking participants' views on the location of an Office of Indigenous Policy Evaluation to oversee Australian Government agencies' implementation of the Indigenous Evaluation Strategy. Which current statutory agency would provide the best location for the Office of Indigenous Policy Evaluation, and why?

An Indigenous governance mechanism

In line with the overarching principle of the Strategy — centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges — and the Australian Government's commitment to 'genuine engagement and partnerships' with Aboriginal and Torres Strait Islander people (chapter 1), oversight of the Strategy should also include governance arrangements that ensure effective engagement with Aboriginal and Torres Strait Islander people.

A number of participants of the Culture, Capability and Governance Roundtable (hosted by the Commission in November 2019 for this project) were of the view that any central body set up under the Strategy should work in partnership with Aboriginal and Torres Strait Islander people. However, several participants also commented on advisory or other consultative bodies often having little influence, with some pointing out that an advisory or consultative role is not the same as partnering or co-design (Tony Dreise et al., sub. 33, p. 3; Lowitja Institute, sub. 50, p. 8; Royal Australian College of Physicians, sub. 77, p. 4).

It is important that any governance arrangements put in place are not tokenistic, but are about genuine engagement and partnership. This will be essential for: creating direct accountability to Aboriginal and Torres Strait Islander people (Plowman 2008, p. 6); developing reciprocity; and centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges.

An Indigenous Evaluation Council (the Council) could be established to provide strategic guidance to the OIPE on its stewardship and monitoring of the Strategy. The Council could also partner with the OIPE to identify whole-of-Australian Government evaluation priorities. In line with the centring principle, the majority of Council members should be Aboriginal and Torres Strait Islander people. The Commission has not yet formed a view on how members of the Council might be appointed, and is seeking participants' views.

INFORMATION REQUEST 7.3

The Commission is seeking participants' views on how members of an Indigenous Evaluation Council might be appointed. For example, could members be nominated by the Aboriginal and Torres Strait Islander community controlled sector; for their experience in research, monitoring and evaluation; or based on some other factors? Would the host agency for the Office of Indigenous Policy Evaluation, and/or the Australian Government, need to be members?

The OIPE and the Indigenous Evaluation Council — how would it work?

The proposed OIPE would be responsible for assessing agencies' performance to ensure their processes and practices align with the Strategy. It would also report publicly on the state of evaluation for policies and programs affecting Aboriginal and Torres Strait Islander people across the APS, including reporting on lessons learned, examples of good (and possibly poor) evaluation practice, potential areas for evaluations on cross-cutting topics, and gaps in the evaluation evidence base. Chapter 10 outlines a proposed approach to assessing agencies' performance against the Strategy, which covers rolling Three Year Evaluation Forward Work Plans and the quality of evaluations whether completed evaluations are made public.

The OIPE could also facilitate the engagement that needs to underpin the priority-setting process for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people (drawing on the evaluation knowledge and experience of organisations both within and outside government: chapter 6).

The Indigenous Evaluation Council would provide strategic guidance to the OIPE on the direction and focus of its advice to agencies, and would partner with the OIPE in identifying Australian Government-wide evaluation priorities (chapter 6) and developing the periodic report on the state of evaluation for policies and programs affecting Aboriginal and Torres Strait Islander people across the APS (figure 7.1). The Indigenous Evaluation Council would also engage with Aboriginal and Torres Strait Islander people, communities and organisations to facilitate their input into the above functions. The OIPE could provide Secretariat support for the Council.

Summarising the proposed governance arrangements for the Indigenous Evaluation Strategy

Governance arrangements for the Indigenous Evaluation Strategy (the Strategy) need to incentivise agencies to produce useful and high-quality evaluations, enable agencies to access guidance when needed, and ensure Aboriginal and Torres Strait Islander people's engagement in the implementation of the Strategy.

The Commission proposes that an Office of Indigenous Policy Evaluation (OIPE) be established within an independent statutory authority of the Australian Government. The OIPE would oversee the Strategy's implementation and coordinate a whole-of-government approach to evaluating policies and programs affecting Aboriginal and Torres Strait Islander people. Its functions should be to:

- have stewardship of the Strategy
- champion the value of evaluation in policy making and identify potential areas for evaluations or systematic reviews on cross-cutting topics
- provide advice to Australian Government agencies on what the Strategy means in practice
- identify, in partnership with the Indigenous Evaluation Council (see below), evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people
- oversee the development and progress of agencies' Three Year Evaluation Forward Work Plans
- coordinate with the Head of Evaluation Profession on capability building opportunities
- monitor and publicly report on Australian Government agencies' performance against the Strategy and on how effective the Strategy has been in improving the quality and usefulness of evaluations affecting Aboriginal and Torres Strait Islander people.

An Indigenous Evaluation Council (the Council), consisting of a majority of Aboriginal and Torres Strait Islander members, could also be established to work with the OIPE. The functions of the Council should be to:

- Provide strategic guidance to the OIPE on:
 - evaluation planning, commissioning, conduct, publication and use
 - capability building for evaluation and cultural safety
 - the translation, dissemination and synthesis of evaluation findings
 - agencies' compliance with the requirements of the Strategy
 - engaging with Aboriginal and Torres Strait Islander people, communities and organisations to facilitate their input into the above functions.
- In partnership with the OIPE:
 - recommend an Australian Government-wide set of evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people
 - report on the state of Indigenous policy evaluation in the Australian Public Service.

'Rewarding and applauding' innovation and good practice

The success of the Strategy does not only depend on agencies complying with the principles-based evaluation framework and the *Guide to Evaluation under the Indigenous*

Evaluation Strategy (and in fact the Strategy will have failed if agencies are only ticking boxes), but also on the extent to which it helps raise the bar on the quality and usefulness of evaluations, and develops a culture of evaluation and learning across the APS.

An important role for the OIPE would, therefore, be to incentivise agencies to evaluate with the purpose of improving policy and program outcomes for Aboriginal and Torres Strait Islander people. This means acknowledging experimentation and innovation (noting that the answers to many policy problems are not known in this area) and engaging with Aboriginal and Torres Strait Islander people in a way that creates value — that is, collecting information that helps policy makers better understand the impacts of policies and programs (and the information is used to reform policies and ultimately improve lives).

The OIPE's role in championing evaluation, and encouraging a focus on the beneficial outcomes enabled by better evaluation, would be particularly important. For example, rather than simply tracking the numbers of evaluations conducted, the OIPE should be looking at what policies and programs are being evaluated, and whether they align with what Aboriginal and Torres Strait Islander people say are the priority areas for evaluation (that is, whether they contribute to filling gaps in the existing evidence base).

Another important part of the OIPE's oversight role would be the dissemination of learnings from agencies' evaluative activities (including examples of how agencies structure their internal management of evaluation, develop capabilities, and share results), with a focus on using these learnings to promote continuous improvement. Innovations in evaluation, including in cases where they have failed, should be rewarded where it is clear that good practice evaluative standards were followed. Evaluations that provide useful insights for other evaluation could also be highlighted. The Western Australian Government suggested:

... reward-based strategies to improve compliance ... such as annual awards for excellence in evaluation practice. (sub. 74, p. 11)

For example, an evaluation of a mainstream policy may not be able to demonstrate impacts on remote Aboriginal and Torres Strait Islander communities (because of limitations in the data), but the process could identify potential ways of overcoming the data limitations, and this could provide valuable learnings for other agencies.

The OIPE could disseminate learnings through: detailed case studies in its regular reporting as part of its monitoring function under the Strategy (chapter 10); updates to the *Guide to Evaluation under the Indigenous Evaluation Strategy*; and contributing to existing communities of practice and other evaluation networks.

A higher performance bar for some, but not all, evaluations?

The Strategy and the performance monitoring arrangements discussed above would only apply to evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. What this means is that agencies undertaking evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people would have their evaluation

performance monitored, while other agencies would not. This could be viewed as applying a higher performance bar for policies and programs affecting Aboriginal and Torres Strait Islander people than for other policies and programs. On the other hand, it also provides an opportunity to demonstrate the benefits of greater oversight and accountability and what that can mean for the quality and use of evaluations.

The question is: if there are benefits from external oversight, why not have oversight for other evaluations undertaken by Australian Government agencies?

Changes are underway to improve APS-wide leadership and embed a whole-of-government approach to evaluation ...

As discussed in chapter 2, in response to recommendations from the Independent Review of the APS, the Australian Government has given the Department of Finance a new central evaluation enabling role. The Department of Finance is to create a small team to ‘develop guidance to ensure systematic evaluation of programs and policies’, and to work with the Department of the Prime Minister and Cabinet to ‘embed evaluation planning in new policy proposals’. Along with the creation of an evaluation profession, agencies will also be required to ‘review and boost their own evaluation capabilities’ under this approach (DPMC 2019d, p. 22).

While there is little detail about the proposed approach, the Government’s response suggests a preference for incremental steps, which largely build on existing efforts and practices across agencies. The question is whether these changes will be sufficient to incentivise long-term changes in behaviour, or to build lasting capacity for evaluation across the APS. Further, due to APS priorities and organisational needs changing in response to the COVID-19 pandemic, the reforms stemming from the Independent Review of the APS have been paused (Gaetjens and Woolcott 2020).

... but key components are missing

The Governments’ response to establishing a centralised approach to evaluation across the APS is missing some key components. They include:

- external oversight or an evaluation auditor to assess the quality and use of evaluations (for stronger accountability, recognising that agencies’ performance is already monitored to some extent, including via Senate estimates and public scrutiny) — the Department of Finance’s central evaluation role appears not to include assessing the quality of evaluations against the evaluation guidelines or how evaluation findings are used. That said, the Independent Review of the APS does refer to a need for the central evaluation function to ‘uphold minimum standards of evaluation’ (2019e, p. 222)
- an evaluation priority-setting process at the whole-of-government level (although this may emerge from the Secretaries’ Board work program for deep-dive evaluations, and

the systematic approach to evaluation, that the Australian Government set out in its response to the Independent Review of the APS: DPMC 2019d, pp. 17, 22, 23)

- a central clearinghouse to store and accumulate the knowledge from evaluation — the Department of Finance has not been asked to accumulate the knowledge from evaluations, or make sense of evaluation learnings across the APS. As discussed earlier, the collection, dissemination and translation of evaluation findings will be critical if the value of evaluation is to be optimised.

A more centralised approach?

Gray and Bray, in the ANZSOG paper for the Independent Review of the APS, recommended a more structured central evaluation management function:

[An agency should take] an oversight role of evaluation, including the development of a strategic approach to ensure evaluation effort is more appropriately targeted and incorporated into policy proposals, and with departments being accountable to the agency for the evaluation they undertake and manage, including regular reporting, and where appropriate, peer review. The agency role should also include reporting on evaluation conduct and results to the government and parliament, along with the public release of evaluations. (Gray and Bray 2019, p. 25)

And Nicholas Gruen³³ has for several years advocated for an ‘Evaluator General’ — an independent statutory authority with responsibility (over and above individual agencies) for all Australian Government evaluations and for encouraging agencies to experiment and innovate more in the design and delivery of policy. Box 7.7 provides more detail on this proposal.

A number of other countries have centralised evaluation oversight arrangements in place (appendix D). Canada’s *Policy on Results*, for example, places a number of obligations on federal departments for performance management and evaluation. The implementation of, and compliance with, the Policy is overseen by the Canadian Treasury Board Secretariat. The Secretariat’s role includes:

- providing leadership for performance measurement and evaluation functions throughout the Canadian government
- reviewing each department’s evaluation plan, and requesting additional evaluations over and above those planned by a department (if necessary)
- initiating centrally-led evaluations (if necessary)
- raising any issues with deputy heads or the President of the Treasury Board (a Minister) about how departments are complying with the Policy
- establishing competencies for the heads of performance measurement and heads of evaluation and amending these competencies as appropriate (Government of Canada 2016b).

³³ Nicholas Gruen is Chair of the Open Knowledge Foundation Australia.

Box 7.7 **An Evaluator-General: Nicholas Gruen's proposal**

The problem: limited mechanisms for 'bottom-up' accountability

Gruen argues that accountability in social policy making needs to flow in two directions: from the top down (as most decision making operates), and from the bottom up (that is, 'validation' from the people delivering, and most affected by, the policy). He also claims that there is a lack of value placed on bottom-up streams of feedback, and this inhibits the drive for innovation coming from people who deliver programs on the ground ('the coalface') who often have a better understanding of what individual communities need.

Gruen contends that the coalface is where experimental policy needs to be developed, and there should be clear avenues for innovative ideas — informed by on-the-ground monitoring and evaluation data — to flow up to senior decision makers.

Those at the centre of the system need to listen to those at the coalface every bit as much as those at the coalface need to listen to those at the centre. After all, those at the coalface are where the action is. ... So the coalface and those at the centre of the system must listen conscientiously to each other to jointly serve the wellbeing of the whole system. (Gruen 2016b)

Gruen also argues that the fidelity and usefulness of evidence are both placed at risk when accountability only flows upwards.

The solution: an Evaluator-General with officers seconded into agencies

To address these issues, Gruen proposes that the Australian Government should create an independent statutory role of an 'Evaluator-General' that would have officers partnered with line agencies' policy and program delivery teams to carry out evaluations, share data and learnings, and encourage greater innovation through bottom-up policy trials.

Evaluation cannot achieve its full potential if it is not assisting those at every level of delivering a program to understand what is being achieved, how, and accordingly, the most promising prospects for improving effectiveness and efficiency. Thus, it must be integrated into service delivery in a way that is collaborative with those delivering services. (Gruen 2018, p. 18)

Evaluation would be done by people with domain skills in *both* evaluation and in the service delivery area who were formally officers of the Office of the Evaluator-General [EG]. They would report to *both* the portfolio agency delivering the program and to the evaluator-general with the EG being the senior partner in the event of irreconcilable disagreement. All data and information gathered would travel to the centre of both the EG's and the departmental systems. Meanwhile, the portfolio agency would report to their minister but the EG would report to Parliament — as the Auditor-General does. (Gruen 2016a)

In 2016, Gruen contemplated that this role could be incorporated into the Australian National Audit Office (ANAO), but doubted that this would provide the right environment for experimentation, arguing that the ANAO's presence may be 'seen as inimical to innovation' (Gruen 2016a). Subsequently, Gruen's submission to the Independent Review of the APS envisaged the Evaluator-General as heading up a new standalone evaluation agency, arguing that building in transparency from the start would enable better collaboration with program delivery teams, sharing of data and learnings, and visibility of evaluation as a profession (needed to highlight its importance in public policy) (2018, pp. 5, 18).

The United States' *Foundations for Evidence-Based Policy Making Act 2018* is another example of a whole-of-government approach to evaluation and evidence building. Similar to the Canadian *Policy on Results*, the Act places several formal obligations on US federal agencies to plan and undertake evaluation. For example, agencies are required to issue an

evaluation plan for each fiscal year that describes the evaluation activities it plans to conduct. These evaluation plans undertaken by agencies are required to describe:

- the key questions that are to be addressed by each evaluation. This may include discussion of a program's purpose, goals and objectives and how program activities are linked to their intended effects
- the methods to be used for evaluations included in the plan, including articulating — to the extent that is practicable — the design of evaluations
- anticipated challenges posed by the evaluations included in the plan (to the extent that this is feasible and appropriate)
- how the agency proposes to disseminate and use the results of evaluation (Executive Office of the President (US) 2019a).

The case for central evaluation leadership and oversight in the Australian Government is wider than the context of policies and programs affecting Aboriginal and Torres Strait Islander people (box 7.8). Longer term, a new independent statutory authority — a Centre for Evaluation Excellence — could be established to provide evaluation leadership, guidance and external oversight for all social and health policy evaluations across Australian Government agencies, and maintain a clearinghouse for evaluation evidence (box 7.9).

If such a Centre was established, the OIPE could move to the Centre as a standalone branch (figure 7.1), with the Indigenous Evaluation Council continuing its role providing strategic direction and partnering with the OIPE's work. The Centre could also have other branches specialising in particular policy areas.

The OIPE could supplement the functions of the Centre by providing specialised advice and guidance on issues specific to the Indigenous Evaluation Strategy (including how Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred throughout the evaluation process). For example, the Centre could provide guidance and advice on the type of evaluation most suitable given the timing or characteristics of a policy or program. The OIPE could advise on whether the method or approach is suitable in light of the cultural or community context, and what other tools could support credible and meaningful evaluations. It could also provide guidance on how the evaluation's results could best be disseminated to the Aboriginal and Torres Strait Islander people, communities and organisations affected by the policy or program.

The Government's response to the Independent Review of the APS — including the new team in the Department of Finance, developing general evaluation guidelines, and a Head of Evaluation Profession — is an important first step towards establishing a Centre for Evaluation Excellence. The Commission suggests that such a Centre be established within five years.

Box 7.8 Why extend evaluation oversight beyond the Indigenous Evaluation Strategy?

Many of the elements of good evaluation practice and governance discussed in this paper are relevant for improving evaluation practice across Australian Government agencies more generally — and in fact, some would be most effective if applied across the board.

As pointed out in earlier chapters, there is significant variation across agencies in terms of the extent to which evaluations are currently undertaken, the quality of those evaluations, and the support that agencies provide for evaluation. Given the direct transferability of skills between better practice evaluations on the whole and those expected under the Indigenous Evaluation Strategy (the Strategy), efforts to embed a broad culture of evaluation and learning across agencies would *also* help support evaluations under the Strategy.

With increased evaluative capacity across all government activities, agencies would become better able to identify where more targeted approaches may be required to support evaluations in scope of the Strategy — particularly for mainstream policies and programs that affect Aboriginal and Torres Strait Islander people. For example, improving the quality of program logic of all government policies and programs would in turn improve the definition and measurement of appropriate outcomes, including for specific population groups (ideally by way of greater engagement *with* those groups). This would help to ensure that a policy's potential effects on Aboriginal and Torres Strait Islander people are considered more consistently, and if necessary, explored more closely under the Strategy. By contrast, an approach that only targets policies and programs judged to affect Aboriginal and Torres Strait Islander people could see key mainstream policies and programs falling through the cracks.

Box 7.9 What functions could a Centre for Evaluation Excellence perform?

A Centre of Evaluation Excellence could:

- advocate for the enhanced generation and use of higher-quality evaluation evidence across the Australian Public Service (APS)
- support agencies' evaluation functions by providing evaluation guidance material and advice, and by coordinating with the Head of Evaluation Profession to deliver evaluation capability building opportunities
- lead the process for identifying evaluation priorities across the whole of the Australian Government, and coordinate inter-agency cooperation on evaluation efforts where needed
- oversee agencies' development and publication of, and adherence to, evaluation forward work plans
- share lessons from evaluation, including by maintaining an online public clearinghouse for evaluation plans, tenders, and published evaluation reports, and by coordinating knowledge translation efforts in order to incorporate evaluation findings into the existing evidence base
- monitor agencies' adherence to good evaluation practice and to the requirements of whole-of-Australian Government evaluation strategies, including the Indigenous Evaluation Strategy
- report every two years on the state of evaluation across the APS.

Figure 7.1 Key functions of the Centre for Evaluation Excellence

Centre for Evaluation Excellence	
	Office of Indigenous Policy Evaluation
Leadership/advocacy: Evaluation champion	Evaluation champion
Capacity building: • Guidance material and advice to agencies on good practice evaluation governance and standards • Collaborate with Head of Evaluation Profession (on skill development, communities of practice, etc.)	• Guidance and advice on Indigenous Evaluation Strategy governance and standards, and good practice • Cultural capability (so evaluations centre Indigenous people, perspectives, priorities and knowledges) • Collaborate with APS Head of Evaluation Profession
Coordinate evaluation effort: Advise on evaluation priority-setting if needed	Government-wide priority-setting under the Strategy
Oversight: Agency evaluation plans	Oversee agency Three-Year Forward Work Plans under the Strategy
Clearinghouse for evaluation reports: Translate and disseminate learnings	Translate and disseminate learnings under the Strategy
Monitor and report: Whole-of-government evaluation performance	Monitor and report on performance under the Strategy

A role for mandatory evaluation requirements?

Mandatory requirements can also be used to embed evaluation into the policy planning and policy making process.

A number of jurisdictions place mandatory evaluation requirements on agencies. For example:

- in Canada, all federal Government departments are required to publish a rolling five-year evaluation plan (covering planned evaluations for all ongoing programs of CAD \$5 million or more annually: appendix D)
- the Western Australian Government operates a similar system — all programs with an annual cost of at least \$5 million have a ‘Sunset Clause’ applied, which means that funding is allocated for a finite period (typically three years) after which a program evaluation must take place to inform the Government’s consideration of funding continuation. New Business Cases for programs over the \$5 million threshold must also include a Treasury-assessed evaluation plan which specifies the program’s desired outcomes and the indicators intended to measure them (Department of Treasury (WA) 2015, p. 5).

As chapter 2 discussed, in the late 1980s and early 1990s the Australian Government had formal requirements for evaluation in place, including that every program was to be evaluated every three to five years; annual portfolio evaluation plans were to be submitted to the Department of Finance, and New Policy Proposals (NPPs) were to include a statement of proposed future evaluation.

The Independent Review of the APS did not explicitly state whether or not the Panel was in favour of bringing back a system of mandated evaluation planning and activity, but did propose that the Department of Finance’s ‘APS-wide approach’ should ‘ensure systematic evaluation of programs and policies’, and that the Department of Finance and Department of the Prime Minister and Cabinet should work together to ‘ensure agencies meet new Cabinet and budget process requirements’ underpinning ‘a systematic approach for formal evaluations’ (DPMC 2019e, p. 223).

As chapter 5 discussed, it is crucial to an evaluation’s quality and usefulness that planning for monitoring and evaluation takes place early in the policy or program design process. There was strong support from participants for New Policy Proposals³⁴ to include evaluation plans, and to demonstrate that a policy’s impact on Aboriginal and Torres Strait Islander people had been considered. This is important in the context of priority-setting being an ongoing process (chapter 6), meaning that a program may not initially (at the design stage) appear to be a high priority for evaluation under the Strategy but may in practice turn out to effect change in a Closing the Gap priority area, or otherwise have a significant impact on Aboriginal and Torres Strait Islander people.

³⁴ It is important to note that evaluation plans in NPPs only deal with the flow of new policies and programs, not the large stock of existing policies and programs (chapter 6 provides more detail).

The Commission agrees that formal policy and program design processes need to have a basic ‘threshold’ analysis of evaluation needs and approaches (see, for example, the components laid out in Coglianese 2012, p. 51), and consideration of a policy or program’s effect on Aboriginal and Torres Strait Islander people, built in.

Embedding evaluation into policy design

One way to embed evaluation into the early stages of the policy design process is to amend Cabinet processes and/or documentation — particularly for New Policy Proposals (NPPs).

Ideally, all NPPs would be required to formally consider how a policy or program might best be monitored and evaluated. However, for the purposes of the Indigenous Evaluation Strategy specifically, the Cabinet Handbook and/or the New Policy Proposal template could be amended to stipulate that:

1. An NPP should state whether the proposed initiative would have a significant and/or differential effect on Aboriginal and Torres Strait Islander people.
2. If so, the NPP should include an Indigenous Evaluation Threshold Assessment (IETA).
3. An IETA 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.

There was also support from participants for evaluation forward work plans to be mandated and for requiring agency management to publicly respond to evaluation findings. However, there was limited support for mandating evaluations for *all* policies and programs, with many participants expressing concerns that this would result in evaluation becoming a ‘tick and flick’ compliance exercise.

Mandating agencies to publicly report their progress against their rolling Three Year Evaluation Forward Work Plans (for example, in their annual reports) could create stronger incentives for agencies to adhere to their Work Plans. It would also create stronger lines of accountability to Aboriginal and Torres Strait Islander people and to the general public (box 7.10).

Ensuring that evaluation findings feed into the budgetary cycle

For the Strategy to achieve its objective, the increased accessibility and usefulness of evaluation results (section 7.2) needs to be complemented by agencies having stronger incentives to use those results when developing and implementing future policies and programs.

Box 7.10 Reporting on evaluation progress

Current arrangements in some agencies lack transparency not only of final evaluation *results*, but also around what evaluations are taking place. Along with Three Year Evaluation Forward Work Plans being published annually, agencies could publicly track the progress of their evaluation activities. This increased transparency would not only create stronger accountability, but would help inform researchers, service providers and Aboriginal and Torres Strait Islander organisations of which gaps in the evidence base were on track to be addressed.

One example of this system already in place is the National Indigenous Australians Agency's (NIAA) Indigenous Advancement Strategy Evaluation Work Plan. As well as detailing the evaluation projects, reviews and enabling activities planned for the current financial year, each Work Plan also records progress made since the previous financial year (i.e. against the previous Work Plan) and, where planned evaluation activities have *not* commenced or been completed on schedule, explains why (NIAA 2019).

There would be benefits in other Australian Government agencies adopting the NIAA's approach of providing annual updates on the status of evaluations planned under the Strategy and listed in their Three Year Evaluation Forward Work Plans (chapter 6). Agencies could publish these details in their annual reports or incorporate progress tracking directly into their Work Plans. For completed evaluations, details could be provided on how the agency has responded to the evaluation findings according to their management response (discussed in section 7.2). And, ideally, where planned evaluation activities under the Strategy had not commenced or been completed on schedule, an explanation would be provided for why this is the case.

This information should also be provided to the Office of Indigenous Policy Evaluation for the purposes of its monitoring and oversight function.

According to the OECD, incorporating evaluation findings in the budgetary cycle (where expenditure and allocation decisions are informed by evaluation findings) is one of the most commonly used tools across OECD countries for promoting the use of evaluations (OECD 2020a, pp. 123–124).

In the Australian context, however, the Independent Review of the APS commented that:

Different mechanisms are available to the APS to assess the effectiveness of spending and its alignment with the Government's policy objectives. ... However, the *ad hoc* nature of these processes can mean a large number of programs continue year after year with no requirement to be periodically assessed. The review supports a more systematic approach to ensure targeted reviews of expenditure. Different options can meet this objective, including a formal requirement to periodically evaluate spending programs (for example, every 3 to 5 years) or a rolling forward program of functional and efficiency reviews. The review recommends that Finance supports regular spending reviews, and helps coordinate the use of evaluation findings, performance information and spending review outcomes to ensure robust advice to government on the effectiveness of current spending during the budget process. (2019e, p. 270)

Along with the Western Australian Government's 'sunset clause' system discussed above, other Australian jurisdictions also have evaluation built more explicitly into the budget cycle. The New South Wales Government, in a Treasury Circular setting out overarching requirements for program evaluation, states that all program-related submissions to Cabinet

which seek funding are to include details of: when the program was last evaluated (if the submission is in regard to renewed funding for an existing program), including evaluation findings and recommendations and the agency response; when it will next (or first) be evaluated; and when evaluation planning will commence (for both existing and new programs) (NSW Treasury 2018, p. 2). The Northern Territory Government stated in 2019 that it plans to embed evaluation planning and results into the policy cycle via a similar system (NT Government 2019, p. 11).

The Office for Development Effectiveness (ODE), within the Department of Foreign Affairs and Trade, is one example of an Australian Government agency that has formal mechanisms to substantively implement the results of its evaluations. Under the 2017 *Aid Evaluation Policy*, all management responses to the ODE's strategic evaluations should be finalised within 28 days of completion of the evaluation report. The completed report, along with the corresponding management response, is then published on the Department's website (DFAT 2017a, p. 3). The Department also publishes an annual *Review of Uptake of ODE Recommendations* on its website, which tracks the progress of initiatives agreed in management responses to the recommendations of the ODE's strategic evaluations.

For the APS as a whole, the Department of Finance previously published an annual paper on the use of evaluation in the Budget, but this was discontinued after 1995 (see, for example, Department of Finance 1994b). Several participants³⁵ argued: that a system for tracking the use of evaluation results should be reintroduced; that there should be formal requirements for evaluation findings to be incorporated into future policy and program design; or that Secretaries or Ministers should be required to publicly respond to evaluation conclusions, as is the case in Canada.

The NSW Aboriginal Land Council, for example, argued that:

... one of the Productivity Commission's remits in developing this Indigenous Evaluation Strategy should be to ensure that no new Indigenous programs are developed and implemented without taking into account previous evaluation findings of similar programs ... (sub. 91, p. 2)

Given the realities of the policy-making process, it is not possible in practice to guarantee that evaluation results have a particular effect on the design of any given policy or program (Tune 2010). Also, some evaluations will not be well-suited to generating recommendations that can be measurably implemented. However, requiring agencies to publicly respond to evaluation conclusions (including a plan for incorporating those conclusions into future policy or program design, as discussed in section 7.2 above) would increase agencies' accountability to Aboriginal and Torres Strait Islander people, communities and organisations, and to the Australian community as a whole. Having agencies report on their management response and progress following an evaluation (box 7.10) would also enable future policy makers and the public to more easily track the path of a policy or program through evaluation and subsequent change.

³⁵ See, for example, AIMN (sub. 15, p. 12); Eva Cox (sub. 64, p. 9); NATSIHSC (sub. 104, p. 4); VACCHO (sub. 44, pp. 22–24).

8 Evaluation capability

Key points

- Evaluation capability refers to the ability of an agency to produce and use evaluation to effectively support accountability and learning. It is about embedding evaluative thinking within an agency so that questions about the value of policies and programs are asked, options to improve them are assessed, and evidence is gathered and used to inform decision making.
- Evaluation capability is an important complement to a culture of evaluation and learning — evaluation cannot be done well without evaluation capability, but it is also difficult to build and maintain evaluation capability unless evaluation is valued.
- Evaluation capability and cultural capability also go hand in hand. Both involve a reflective way of doing and thinking and the skill and mindset to understand the impacts of policies and programs on different groups within the community — including Aboriginal and Torres Strait Islander people.
- Evaluation capability varies across the Australian Public Service (APS). Because agencies are at different stages of maturity, some will need to invest in their evaluation capability if they are to comply with the Indigenous Evaluation Strategy (the Strategy).
- Evaluation skills are not just important for evaluators. Those commissioning evaluations, project managers and APS staff need to be able to understand, plan for and use evaluations.
- Strengthening the capability and increasing the number of Aboriginal and Torres Strait Islander people to undertake or participate in evaluations is also important.
- Agencies will be responsible for ensuring that they have access to the skills they need to undertake or commission evaluations under the Strategy. However, central support should be provided in the form of guidance, training and development.
 - The Strategy and the accompanying guidance support good practice evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.
 - The Head of Evaluation Profession — agreed to by the Australian Government as part of its response to the Independent Review of the APS — should also provide support that agencies can draw on to build and maintain their evaluation capability.
- The decision about whether to use an internal or external evaluator should be based on who is best placed to undertake a particular evaluation. However, 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, including the extent to which Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred.
 - For commissioned evaluations, agencies should report on how their commissioning arrangements have centred Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges when reporting on how they have upheld the principles of the Strategy.

Evaluation capability is important for delivering high quality evaluations. It is the ability of agencies to produce and use evaluation to effectively support accountability and learning. It is about agencies asking the right evaluation questions; having the ability to collect, analyse, and interpret data; and using evaluation findings for decision making and action. It means that agencies are able to effectively plan and commission evaluations.

Evaluation capability also includes cultural capability. As the Aotearoa New Zealand Evaluation Association said:

Cultural competency is central to the **validity** of evaluative conclusions as well as to the appropriateness of the evaluation process. Cultural competency in evaluation goes beyond conducting evaluations in culturally appropriate, responsive or sensitive ways. It also means drawing on the values, needs, strengths and aspirations of the culture of those a policy or programme is intended to benefit to define what is meant by ‘good program content and design’, ‘high quality implementation and delivery’ and ‘outcomes of value’. (2011, p. 10, emphasis in original)

Evaluation capability is not only important for conducting high quality evaluations, it can also increase the demand for good evaluation and equip agencies with the skills to use evaluation findings effectively. As Bray, Gray and ‘t Hart noted:

... beyond those working on evaluation, attention also need[s] to be given to the wider development of the evaluation literacy of public servants across the [Australian Public Service], so they can more effectively use the products of evaluations, and indeed the development of data and related interpretation skills to best use these sources in policy and program development. (2019, p. 18)

This chapter looks at the evaluation capability that will be needed to support the Indigenous Evaluation Strategy (the Strategy). The chapter covers:

- the current state of evaluation capability across the Australian Public Service (APS) (section 8.1)
- the evaluation and cultural skills required by those undertaking evaluation (section 8.2)
- commissioning high-quality evaluations in a way that reflects the principles of the Strategy (section 8.3).

8.1 What is the current state of evaluation capability in the APS?

Evaluation capability varies across the APS. As discussed in chapter 2, some agencies have dedicated evaluation units and evaluators. Many departments also manage evaluation procurement panels to contract out evaluation services, such as the Research, Evaluation and Data Panel (maintained by the Department of Social Services) and the Collaborative Evaluation, Research and Planning Panel (maintained by the National Indigenous Australians Agency (NIAA)). However, a number of agencies told the Commission that they had few staff

with expertise in evaluation and/or the capability to effectively commission evaluations (or to assess the quality of commissioned evaluations).

The NIAA said that ‘significant improvements in evaluation capability and capacity are needed’ and:

... agencies are at different stages of maturity in developing frameworks and undertaking evaluation. A challenge we all need to be aware of is that there may be key ingredients missing to support, enable, and resource evaluations. (sub. 86, p. 4)

The Australasian³⁶ Evaluation Society, in its submission to the Independent Review of the APS, also questioned whether the APS had sufficient in-house expertise in evaluation:

There is ... a broader question as to whether there is a sufficient critical mass of APS staff with expertise in research, evaluation and performance measurement. This is not a recent question — senior public servants have previously expressed a view that key skill sets in research and analysis and evaluation within the APS are in short supply. This may reflect both a capability and capacity issue, possibly arising from the trend for several decades ... for staff to be ‘generalists’ who are capable of and can be deployed to undertake the wide range of tasks often required of public servants — while specialist expertise is often purchased or procured. (2018, p. 10)

And a research paper undertaken for the Independent Review of the APS suggested that issues of workforce capability were secondary to organisational and cultural issues:

... the major impediments to evaluation capacity and capability in the APS are cultural and organisational rather than a lack of skills or analytical capacity. (Bray, Gray and ‘t Hart 2019, p. 18)

The Independent Review of the APS concluded that ‘APS in-house research and evaluation capabilities and processes have fallen’ and called for the Australian Government to ‘rebuild APS research and evaluation expertise’ (DPMC 2019e, pp. 181, 221).

A number of participants also said that there was scope to improve the cultural capability of evaluations (box 8.1).

³⁶ The Australasian Evaluation Society is now known as the Australian Evaluation Society.

Box 8.1 **Participants said cultural capability could be improved**

Social Ventures Australia:

We have observed gaps in knowledge and expertise from evaluation commissioners when considering the scope and methods required for an evaluation to produce a high quality and rigorous product. ... when you're not collecting data in a culturally capable manner then the data you are collecting will be of a poor quality. (sub. 83, p. 15)

National Aboriginal Community Controlled Health Organisation:

... the extent to which cultural competency and capability are addressed in current evaluation practice varies considerably. Only around half of Australian Government and state/territory health program evaluation reports commissioned between 2007 and 2017 integrated aspects of 'cultural respect'. (sub. 95, p. 5)

National Social Security Rights Network:

Our members and ally organisations have observed that evaluation indicators often do not address relevant key questions about particular programs, and evaluators often lack cultural competency and awareness when working with communities. (sub. 10, p. 3)

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) — while also raising concerns about consultants' cultural capability and their understanding of ethical practices when undertaking evaluation — acknowledged that there had been some improvement in recent years.

Where evaluations were carried out, they were most often at arm's length and independent of the Indigenous organisations and communities themselves ... They were often undertaken by consultants that did not have the cultural competency and ethical foundations for the work. In recent years, the maturing of the consulting sector in relation to Indigenous cultural competency and ethical research, together with the Indigenous procurement strategy have seen a shift in practice, although there is a significant way to go. (sub. 72, p. 4)

As discussed in chapter 2, agencies are at different stages of maturity in both their evaluation culture and capability. Some agencies will need to invest in their evaluation capability if they are to implement the Strategy effectively.

8.2 **The skills required to undertake and commission evaluations**

Evaluation capability needs to exist at an organisational level, so that evaluation is not seen as a last minute add on, but rather as something that adds value to the day-to-day business of an agency. It also needs to exist within a culture of evaluation and learning, as without such a culture, evaluation capability is unlikely to be built and sustained over time. In

highlighting the importance of building organisational capability, Williams talked about the risk of being ‘all dressed up with nowhere to go’:

All the skills, knowledge, technical expertise and experience in the world won’t help ... if the ... program, community, organization or environment cannot sustain and nurture those skills, and abilities. (2001, p. 1)

What this means is that evaluation capability needs to be built up across those undertaking evaluation (evaluators), those commissioning evaluations (program managers/policy makers) and those participating in evaluations (such as service delivery organisations and service users who often provide critical information to inform evaluation). Agencies also need to understand the purpose and benefit of conducting evaluations and how to use evaluation findings to improve the effectiveness of policies.

What skills are required for evaluation?

For the Strategy to make a difference, evaluative thinking — a form of critical thinking where evaluative questions are asked as a matter of habit (box 8.2) — needs to be embedded throughout agencies.

Not just evaluators need evaluation skills

The skills required to effectively support evaluative thinking will vary depending on an individual’s role in an evaluation. A person commissioning an evaluation, for example, requires fewer technical skills than an evaluator. That said, as Gray and Bray noted, evaluation skills are also important for those commissioning evaluations.

Enabling an effective evaluation function, even where this is more focused on the management of externally commissioned evaluation functions, requires staff with a suite of technical skills. These include in the area of modelling, critical to being able to undertake *ex ante* evaluation of policy options, and strong data and conceptual skills and a knowledge of evaluation approaches ... (2019, p. 25)

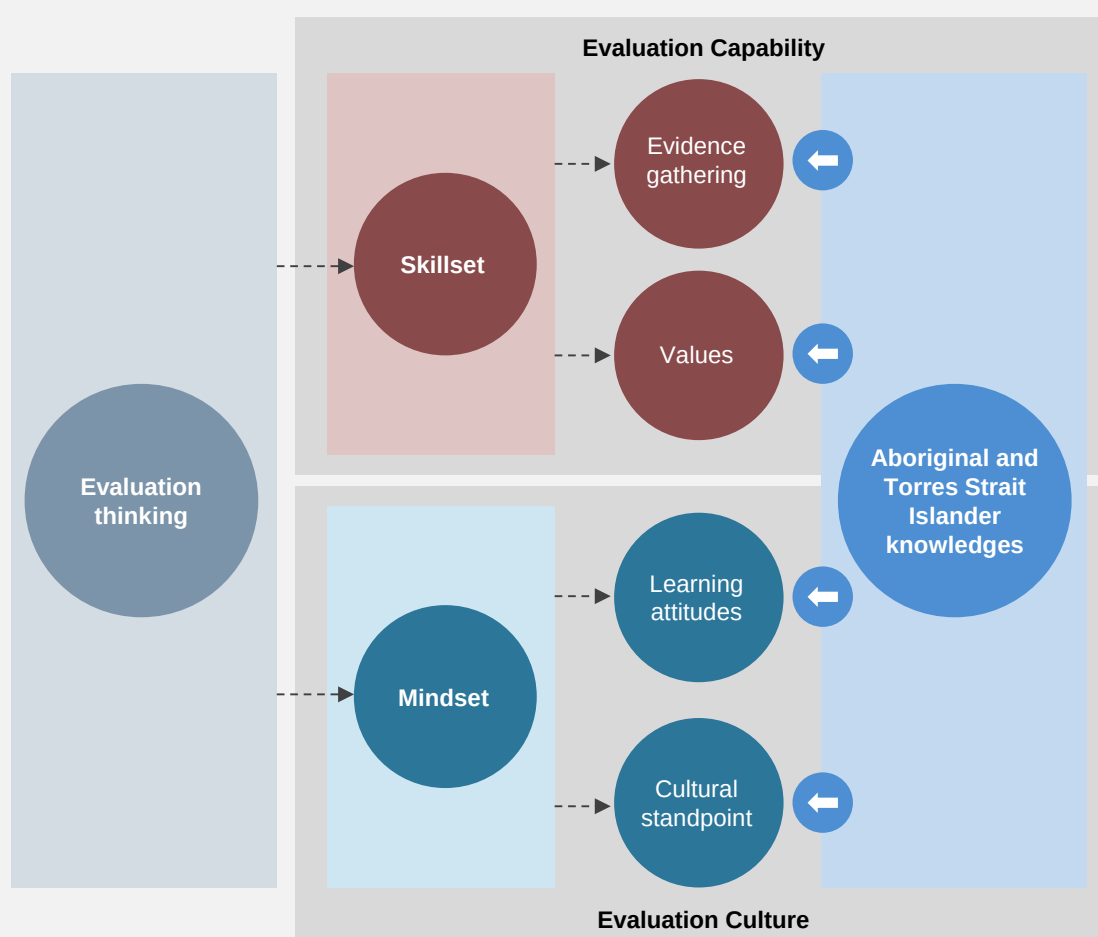
Commissioners of evaluations need to know what competencies are required when selecting evaluators, be able to manage evaluations effectively and assess the quality of an evaluation when it is completed. Commissioners also need to understand what culturally capable, ethical evaluation looks like.

Box 8.2 Evaluative thinking for the Indigenous Evaluation Strategy

Evaluative thinking is a form of critical thinking where, as part of everyday business, evaluative questions — that is, questions about the quality and value of something — are asked to inform evidence-based decision making. Evaluative thinking is part mindset and part skillset. It requires an open and curious mindset, supported by a culture of evaluation and complemented by the skills required to collect and assess the evidence needed to answer evaluative questions effectively.

The skillset needed to support evaluative thinking is one of evaluative reasoning — where quality and value are defined and evidence is collected against those criteria. In the context of evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, evidence gathering should be appropriate to the cultural context. This skillset needs to be paired with value judgements that reflect what ‘quality and value’ means for Aboriginal and Torres Strait Islander people and what evidence is necessary to support those judgements.

Mindset is influenced by an agency’s evaluation culture and involves being curious and open to learning about what works best for Aboriginal and Torres Strait Islander people. This mindset needs to be paired with a recognition of, and willingness to learn about one’s own cultural biases.



Given that all elements of evaluative thinking require recognising the place of Aboriginal and Torres Strait Islander knowledges, cultural capability should not be divorced from investments in evaluative thinking. Cultural capability will influence how the skillset and mindset needed for evaluative thinking are applied.

Sources: adapted from NSW Department of Education (2020); Wehipeihana and McKegg (2018).

A number of participants pointed to the importance of improving project managers and staff's evaluation competencies³⁷. Agency staff should be able to:

- understand what type of questions an evaluation could answer and what information is needed to answer them
- plan for an evaluation (for example, by being able to articulate a program logic and identify potential data sources that could be used to answer evaluative questions)
- collect and interpret monitoring data
- interpret and incorporate evaluation findings into policies or programs.

Undertaking evaluation requires a range of competencies ...

Evaluation is not a profession. There is no recognised set of qualifications for evaluators, evaluators can have training and experience in a range of disciplines, such as economics, medicine, education and public policy. As Rossi, Lipsey and Henry said:

Much discourse has occurred about evaluator competencies ... but it has yet to be codified into a recognized set of qualifications required for anyone conducting evaluation. It remains accurate to describe evaluators as a collection of individuals sharing a common label, who are not formally organized, and who may have little in common with one another in terms of the range of activities they undertake or their approaches to evaluation, competencies, organizations within which they work, and perspectives. (2018, p. 300)

Evaluators need to be able to apply systematic methods of inquiry that are appropriate to the evaluation in question to ensure credible results. Evaluators must also be culturally capable and have the skills to conduct evaluations ethically.

Evaluation requires a mix of technical, management and communication skills. Some evaluation associations have established lists of evaluator competencies (figure 8.1) that cover the skills, knowledge and experiences evaluators should have. Common competencies relate to: the ability to undertake systemic inquiry; technical, reflective, interpersonal and project management skills; and the ability to undertake contextual analysis and engage effectively.

³⁷ See for example, Darling Downs and West Moreton Primary Health Network (sub. 6, p. 3), Closing the Gap Steering Committee (sub. 71, p. 6) and NACCHO (sub. 95, p. 5).

Figure 8.1 Types of evaluation competencies
A selection of different evaluator skill competency frameworks

	AES ^a	UN ^b Evaluation Group	ANZEA ^c	Canadian Evaluation Society
Technical	Evaluation activities	Evaluation skills	Systematic evaluative enquiry	Technical practice
	Evaluation theory			
	Research methods and systematic inquiry			
Reflective		Professional foundations	Reflective practice and professional development	Reflective practice
Contextual	Culture, stakeholders and context		Contextual analysis and engagement	Situational practice
			Values and cultural competency	
Management	Evaluative attitude and professional practice	Promoting a culture of learning for evaluation	Evaluation project management and professional practice	Management Practice
	Project management	Management skills		
Interpersonal	Interpersonal skills	Interpersonal skills		Interpersonal practice

^aAustralian Evaluation Society ^bUnited Nations. ^cAotearoa New Zealand Evaluation Association.

Sources: AES (2013a); ANZEA (2011); CES (2018); UNEG (2016).

There are also skills, knowledge and attitudes that are important for undertaking evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. The Australian Public Service Commission outlines some key competencies that APS employees

should possess when undertaking evaluation and research involving Aboriginal and Torres Strait Islander people. These include:

- understanding ethical research guidelines; good practice research methods; the heritage, cultural and intellectual property rights of Aboriginal and Torres Strait Islander people as they apply to research; that research should be for the benefit of Aboriginal and Torres Strait Islander people; and the impact of mainstream research on Aboriginal and Torres Strait Islander people
- applying evidence-based, ethical and culturally acceptable research practices; ensuring equitable benefits for participating communities; upholding community expectations and cultural protocols; and gathering, interpreting, analysing and presenting data in ways that are most useful and least disruptive for Aboriginal and Torres Strait Islander people
- committing to meaningful engagement and reciprocity; identifying verbal and non-verbal cues; being empathetic and non-judgemental; showing respect for Aboriginal and Torres Strait Islander cultures, knowledges and materials; and engaging and negotiating in good faith (2015, p. 16).

... complemented by cultural capability

Cultural capability is a reflective way of doing and thinking (or an attitude towards culture) rather than a particular set of skills or knowledge. Cultural capability emphasises improvement and adaptation to ‘varied and changing circumstances and responsibilities’ (Russell-Mundine 2017, p. 44). The BetterEvaluation website described a culturally competent evaluator as:

... one who is ready to engage with various sections of communities to embrace cultural and contextual dimensions important to the evaluation. Critical self reflection could be considered the first building block towards accomplishing cultural competence as it is not only the cultural context of the evaluatee that needs consideration but values, belief and culture of the evaluators are also important. Culturally competent evaluators not only respect the cultures represented in the evaluation but recognize their own ‘culturally based assumptions’; take into account the ‘differing world view of evaluation stakeholders and target communities’ and select culturally appropriate evaluation options and strategies. (2016)

Cultural capability is important to ensure the validity of evaluation results. According to the BetterEvaluation website:

Without recognizing the influence of culture on human behaviour, the evaluators may arrive at conclusions which are fundamentally flawed. (2016)

In a similar vein, Social Ventures Australia noted that if data are not collected in a culturally capable manner then the data ‘will be poor quality and of little value’ (sub. 83, p. 10).

It is important that evaluators ensure they have sufficient cultural capability before undertaking evaluations involving Aboriginal and Torres Strait Islander people. An important part of an evaluator’s professional development is increasing their knowledge of

— and experience in — working with Aboriginal and Torres Strait Islander people. But there are also a number of ways to improve cultural capability through how an evaluation is undertaken. For example, the Western Australian Government suggested that the cultural competency of an evaluator could be enhanced by:

- practitioners and commissioners working directly with Aboriginal and Torres Strait Islander people when developing and assessing evaluation methodologies
- working with Indigenous consultants and trusted partners who can advise on cultural competency or safety
- employing and/or promoting Aboriginal and Torres Strait Islander people as commissioners and practitioners, supported by evaluation training and mentoring
- commissioning external Indigenous evaluators, and/or other evaluators who have expertise in engaging with Aboriginal and Torres Strait Islander people
- documenting and sharing lessons on engaging with Aboriginal and Torres Strait Islander stakeholders in evaluations within and between government agencies (sub. 74, p. 8).

Aboriginal and Torres Strait Islander evaluators

Employing Aboriginal or Torres Strait Islander evaluators can improve the cultural capability of an evaluation. Wehipeihana noted that evaluations led by Indigenous people are more likely to have the cultural capital to:

- facilitate respectful engagement and the observance of cultural protocols; they know what matters to get relationships and the evaluation off to a good start
- use Indigenous knowledges, methods, and ways of working in an evaluation; they can provide a cultural ‘reading’ or assessment of the cultural fit or appropriateness of data collection methods and tools for Indigenous peoples
- facilitate an understanding of what value and goodness look like through Indigenous world views; they can take the lead in sense making and analysis to ensure that the richness, subtlety, and nuance of meaning are not lost in translation and ensure the cultural validity of the evaluative conclusions drawn (2019a, p. 372).

That said, not every evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people will need Aboriginal or Torres Strait Islander evaluators. The decision about who should evaluate will need to take into account both the technical and cultural capability needed for a high-quality evaluation.

Some participants said that more investment in the capability of Aboriginal and Torres Strait Islander people to lead evaluations was needed. For example, the National Family Violence Prevention Legal Services Forum said:

Unless there is a constant and intentional effort to build a pool of Aboriginal and Torres Strait Islander evaluators, or build the capacity of organisations already delivering programs, we will only ever move between evaluations done ‘to’, ‘for’ and at best ‘with’ Aboriginal and Torres

Strait Islander people. We will also be stuck in a constant cycle of educating and resourcing non-Aboriginal and Torres Strait Islander evaluators and commissioners to be ‘culturally safe’. (sub. 66, p. 10)

The Strategy will provide more opportunities for Aboriginal and Torres Strait Islander people to lead evaluations by raising the bar on the quality of evaluations, which in turn will increase the demand for culturally capable evaluation. There will also be opportunities for more Aboriginal and Torres Strait Islander people to gain skills and experiences in other aspects of the evaluation process, such as advising on evaluation design, undertaking data collection and interpreting results. This should strengthen the capability of Aboriginal and Torres Strait Islander people to lead evaluations in the future.

Participants provided other suggestions to strengthen the capability of Aboriginal and Torres Strait Islander people. Suggestions³⁸ included:

- recruiting and developing Aboriginal and Torres Strait Islander APS employees to be evaluators
- funding and supporting qualifications or training for Aboriginal and Torres Strait Islander people to deepen their evaluation skills, for example, through scholarships to do a Masters of Evaluation or school-based traineeships
- supporting community controlled organisations to strengthen their evaluation capability
- employing and training community researchers to undertake parts of an evaluation
- improving how evaluations are commissioned to better support Indigenous-owned businesses.

Agencies should be responsible for meeting their own skill needs ...

Agencies will need to ensure that they have access to the skills and competencies needed to comply with the Strategy. The composition of skills that agencies require will vary (box 8.3), and will depend on:

- the size, scope and nature of the policies and programs affecting Aboriginal and Torres Strait Islander people that they administer
- the nature of the evaluations they expect to undertake under the Strategy (that is — the type of evaluations they expect to undertake and/or whether they intend to commission evaluations or undertake them in-house).

³⁸ See for example Central Australia Academic Health Science Network (sub. 85, pp. 2-3), Maggie Walter (sub. 112, p. 2), NSW Aboriginal Land Council (sub. 91, p. 6) and University of Queensland (sub. 20, pp. 6-8).

Box 8.3 **Capability building — the experience of two Victorian government departments**

The Victorian Department of Health and Human Services (DHHS) and the Victorian Department of Education and Training (DET), while both having a central evaluation unit, have adopted different approaches to building capability.

Both evaluation units provide training, guidance, knowledge translation and advisory services. But the evaluation unit of the DHHS focuses on undertaking evaluations themselves, whereas the evaluation unit of the DET focuses more on providing advice to line areas and knowledge translation or evidence synthesis activities. Consequently, the size of each unit is considerably different — the DHHS employs about 20 evaluators, whereas the DET employs only a few evaluators.

Source: Williams (2019).

Given the wide variation in the depth of existing evaluation skills and experience across agencies, some agencies will need to devote considerable resources to building their evaluation skillsets and/or their ability to undertake evaluation in a culturally capable manner if they are to implement the Strategy effectively. There are a number of ways that this may be done.

- *Recruiting individuals to address skill gaps* — recruitment provides a way to quickly build expertise, however, it does not lift the skills of the whole department and could leave evaluation or cultural expertise in small, isolated pockets.
- *Providing training and development opportunities* — training and development is a longer-term activity that can grow the pool of evaluators and lift the skills of project managers and evaluation commissioners. Several agencies told the Commission that they provide internal training for staff to strengthen their evaluation skills. There are also many courses and training programs (box 8.4) — both within and external to the APS that focus on improving evaluation skills. Several agencies noted that bespoke training, which focuses on a specific policy or program, is useful to demonstrate the value of evaluation and to develop evaluation skills.
- *Learning by doing and other experiential opportunities* — secondments across the APS can give staff opportunities to develop their evaluation skills. Working directly with Aboriginal and Torres Strait Islander people and communities is another way to build cultural capability.
- *Providing guidance and support for staff to undertake evaluation* — this can take many forms, including: creating good practice evaluation guides, setting up structures that support evaluation (such as maintaining a central team of evaluators to provide support), providing opportunities for formal or informal coaching, and encouraging staff to utilise evaluation networks or communities of practice.

Box 8.4 **Training and development for evaluation and cultural capability**

A number of organisations provide training and development opportunities that agencies could use to improve the evaluation and cultural capability of their staff. These include:

- the Australian Evaluation Society — provides training and development opportunities throughout the year. The Society also runs an annual conference (and usually has sessions on evaluating policies and programs that affect Aboriginal and Torres Strait Islander people) and several regional networks and communities of interest
- the Australian Institute of Aboriginal and Torres Strait Islander Studies — the Institute's CORE program is an online cultural capability course. The Institute also puts on a variety of symposiums, seminars and conferences that cover research, evaluation and engagement with Aboriginal and Torres Strait Islander people
- the University of Melbourne — offers online courses in evaluation, including a graduate diploma and masters course
- Jawun program — offers Australian Public Service (APS) employees secondment opportunities to work in Aboriginal and Torres Strait Islander organisations
- the Institute of Public Administration Australia — runs once-off events and training on evaluation
- Australia and New Zealand School of Government — runs once-off events and training on evaluation. It also convenes the annual First Peoples' Public Administration Conference
- Aboriginal and Torres Strait Islander Community Controlled Organisations — many offer cultural capability training and have developed useful resources. They have specialised expertise in engaging in particular policy areas or geographies.

There are also several communities of practice — such as the APS-wide evaluation community of practice — that provide opportunities for staff to further develop their capability.

Agencies will be best placed to determine what skills they need and the best way to source or develop these skills. Placing the onus for skill building on agencies means that they will have the freedom to invest in the skill development activities that provide them the greatest benefit based on their needs.

... with some broad central support

Notwithstanding that ultimate responsibility for developing and maintaining evaluation skills should lie with agencies, a central function should also provide evaluation support and resources that agencies can access to build or maintain evaluation capability. This function would also allow for APS-wide learning and the sharing of lessons learnt across agencies.

The Strategy itself is one form of central support. *A Guide to Evaluation under the Indigenous Evaluation Strategy* provides initial guidance on how the principles should be applied in practice, as well as questions that agencies and evaluators should consider at each stage of the evaluation process. As part of the broader monitoring process outlined in

chapter 10, the *Guide* should be reviewed and updated so it remains contemporary and fit-for-purpose.

In response to the Independent Review of the APS, the Australian Government has also indicated that the Department of Finance will establish a small team to help build evaluation expertise and practice. The Department will also develop ‘guidance to ensure systematic evaluation of policies and programs in line with the Enhanced Performance Reporting Framework’ (DPMC 2019d, p. 22).

As part of the professions model outlined in the Independent Review of the APS (box 8.5), the Australian Government also agreed to establish an evaluation profession.

Box 8.5 The Independent Review of the APS professions model

The *Independent Review of the Australian Public Service (APS)* recommended that the Australian Government establish an APS professions model for core delivery, regulatory and policy roles, enabling functions and specialist areas. It recommended that each profession:

- be led by a head of profession
- have a set of core competencies and a learning and development framework established by the head of profession (with support from the Australian Public Service Commission)
- develop specialist career pathways within the APS.

The Review recommended that one of the professions established be an evaluation profession. The Australian Government agreed to this recommendation.

The first profession, the human resources profession, was launched in October 2019 (prior to the Review’s release). A head of profession and a professional network have been established, and it has been announced a centralised graduate recruitment process will take place for this profession for the 2021 graduate intake. Further work is underway to establish professional competencies and to undertake a skills audit for the human resources profession, as well as establishing a mobility program to provide secondment opportunities.

Sources: APSC (2019, 2020a, 2020b); DPMC (2019e, p. 200, 223, 2019d, pp. 20, 22).

While the specific approach that the evaluation profession will take is yet to be established, if it is consistent with the recommendation of the Independent Review of the APS, part of its remit will include developing core competencies for evaluators and establishing a learning and development framework to build evaluator capabilities.

This means that the Head of Evaluation Profession is likely to be well placed to ‘build in’ capabilities for evaluators relevant to the Strategy (and development opportunities to improve these capabilities) while developing the competency and learning and development frameworks for the evaluation profession as a whole.

Integrating capability development for the Strategy with capability development for the evaluation profession has two key advantages. First, it minimises the risk of duplication. If capability development activities for the Strategy are planned and organised separately to

development activities for the profession more broadly, there is a risk that these activities will be duplicative or poorly integrated. Second, housing capability development for the Strategy within the evaluation profession sends a signal that the ability to undertake effective evaluations under the Strategy (and, by proxy, the ability to undertake effective and culturally competent evaluations involving Aboriginal and Torres Strait Islander people) is a key capability of all evaluators. The Strategy should be viewed as a key part of — not separate or supplementary to — evaluation in the APS more broadly.

By way of specific roles, the Commission considers that the Head of Evaluation Profession is well placed to:

- provide training for commissioning, conducting, and using evaluations of policies and programs that affect Aboriginal and Torres Strait Islander people
- facilitate an APS-wide community of practice for people who are involved in evaluating policies and programs that affect Aboriginal and Torres Strait Islander people
- establish processes through which evaluators can seek secondments or other mobility opportunities, with a view to broadening or deepening their experiences evaluating policies and programs affecting Aboriginal and Torres Strait Islander people.

In developing core competencies for evaluators and establishing a learning and development framework, the Head of Evaluation Profession should consider:

- the skills and competencies needed to effectively conduct evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people
- the skills and capabilities needed by different parties — including policy makers, program managers, evaluation commissioners and evaluators — to undertake and use evaluations effectively
- how existing expertise on evaluation involving Aboriginal and Torres Strait Islander people should best be used — this should include the expertise of AIATSIS, the Lowitja Institute and Aboriginal and Torres Strait Islander organisations, researchers and academics.

A central approach to growing a cohort of Aboriginal and Torres Strait Islander evaluators in the APS would help to implement the Strategy. As identified above, Aboriginal and Torres Strait Islander evaluators may improve the cultural capability of an evaluation. A network of Aboriginal and Torres Strait Islander evaluators may also help shift the evaluation culture in the APS to focus on outcomes that are important to Aboriginal and Torres Strait Islander people and help further agencies' understanding of Aboriginal and Torres Strait Islander research methods. Ewen et al. found similar benefits in health research workforces, they found that peer networks:

... enable opportunities for researchers to combine shared experiences and identities, to grow understandings of research, and to build efforts for a collective push to impact on, Aboriginal and Torres Strait Islander health outcomes. (2019, p. 30)

The Head of Evaluation Profession is well placed to develop a strategy to build a cohort of Aboriginal and Torres Strait Islander evaluators within the APS. As part of developing this strategy, the Head of Evaluation Profession should consider the value of a range of actions, including providing structured training, networking opportunities, mentorship programs and secondment opportunities.

Monitoring the Strategy will provide an opportunity to assess agencies' efforts to improve their evaluation skills

Even with clearer expectations and access to support, there remains a risk that without the right incentives, agencies could underinvest in evaluation skills.

An important part of monitoring the Strategy (chapter 10) will be examining agencies' efforts to develop and maintain evaluation skills to meet their responsibilities under the Strategy. This monitoring should assess the quality of evaluations, with a view to identifying any evidence of gaps (in evaluation skills or in the cultural capability of evaluators).

The monitoring process will help to lift the capability of the APS as a whole. It will assess the guidance provided through the Strategy and whether further clarification or guidance is needed. It will also act as one mechanism, among others, to identify and share good practice and lessons learnt.

8.3 Capability to commission high quality evaluations

Current commissioning practice

Commissioning organisations external to the APS to undertake evaluations is relatively common practice. An external evaluator was used in about 60 per cent of the evaluations in the sample of evaluations received by the Commission from government agencies (chapter 3).

Many agencies are likely to continue to rely predominately on commissioning evaluations to undertake evaluations under the Strategy, as building in-house expertise to conduct evaluations will take time (and for some smaller agencies, strong in-house evaluation expertise may not be viable).

However, some participants noted issues with current commissioning arrangements. For example, Indigenous Community Volunteers raised concerns about ethical practice:

... the evaluation sector has a range of barriers to overcome. In particular, the need for commissioning processes to consider and mandate ethical practices and allow time and adequate budget for the ethical practices to be adhered to. (sub. 88, p. 17)

The Research School of Population Health (Australian National University) also raised concerns about how Aboriginal and Torres Strait Islander governance operates in commissioning processes:

Our observations are that, in certain instances, processes have not always been transparent particularly in how Indigenous evaluators and/or Indigenous forms of governance are utilised. (sub. 84, p. 3)

Others were concerned with the cultural capability of evaluators in commissioned arrangements:

The Australian Government acts under robust financial accountability and ethical conduct frameworks that govern its procurement and conduct. However, the engagement of independent evaluators in these frameworks and systems may not always support cultural competence. Cultural competence, whereby there is established knowledge of, and relationships with, the Aboriginal and Torres Strait Islander communities involved are key to beneficial stakeholder relationships. (Lowitja Institute, sub. 50, p. 6)

There may also be issues with the way governments tender for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. Kelaher et al. (2018b) reviewed a limited sample of tender documents and evaluation reports for evaluations of policies and programs affecting the health and wellbeing of Aboriginal and Torres Strait Islander people. They found that tender documents and evaluation reports often did not reflect or incorporate Australian, state and territory government agencies' own principles about working with Aboriginal and Torres Strait Islander people (as described in health department planning documents).

Is it better to conduct evaluations internally or externally?

There is no definitive answer on whether evaluations should be conducted internally or externally. There are a number of factors that should be considered when deciding who is better placed to conduct an evaluation.

Evaluators conducting evaluations within an agency are likely to have a better understanding of what a policy is seeking to achieve, the origins of a policy and policy processes than evaluators outside the agency. Internal evaluators may also be better placed to build relationships with stakeholders during the planning and design of an evaluation. According to Schoenefield and Jordan, 'careful attention to such factors may eventually facilitate the uptake of evaluation knowledge later on in the policy process' (2017, p. 277).

The purpose of an evaluation can also influence who should undertake it. An internal evaluation is likely to be more appropriate when the purpose of an evaluation is learning and improving the implementation of a policy or program. In contrast, evaluation by external evaluators may be more appropriate when the purpose is to improve accountability, when seeking a new perspective (or a 'fresh set of eyes') on a policy or when specialised skills are required. Internal evaluators may be less critical of their own policies and/or may look for

evidence to support pre-existing hypotheses or views about a policy (Nickerson 1998; Weiss 1993).

That said, while evaluations conducted outside an agency can be more independent, commissioned evaluations can have similar problems to internal evaluations because evaluation criteria are determined by agencies and there can be an incentive to influence evaluators. A survey of evaluators in the United Kingdom found that government used a range of strategies to influence evaluation outcomes, including setting the questions to be addressed by evaluators and putting in place budgetary-turned-methodological constraints (such as not funding a control group) (Hayward et al. 2014). Pleger and Sager (2018) found that similar strategies were employed in other countries. Giving evaluators a short time to conduct evaluations is another strategy that can lead to superficial results (Schoenefeld and Jordan 2017).

Gray and Bray also observed that ‘external evaluators may be influenced by factors such as wanting to keep departments and ministers satisfied so as to influence possible future contracts’ (2019, p. 20). And Hudson asked ‘how much independence can a consultant claim to have when they are reliant on their clients for business?’ (2017, p. 10) A number of participants³⁹ pointed to the Cashless Debit Card evaluation (box 5.2) as a potential example. This evaluation arrived at conclusions supporting the effectiveness of the trial, despite evidence to the contrary (Hunt 2017).

For these reasons, evaluation transparency and oversight are critical. They provide the means by which evaluations are subjected to wider scrutiny. Gray and Bray commented that:

... issues of independence, expertise and quality are not categorically associated with internal versus external evaluation. ... These elements may also be sensitive to the extent to which evaluations are subject to wider scrutiny, including through the discipline of external review ... (2019, p. 20)

Another model involves some aspects of an evaluation being undertaken externally and other aspects undertaken internally by an agency. For example, an agency may commission an external organisation to undertake data collection, but analyse and synthesise these data in-house. Alternatively, an agency may commission one external organisation to develop a plan for an evaluation and then commission a different organisation to undertake it.

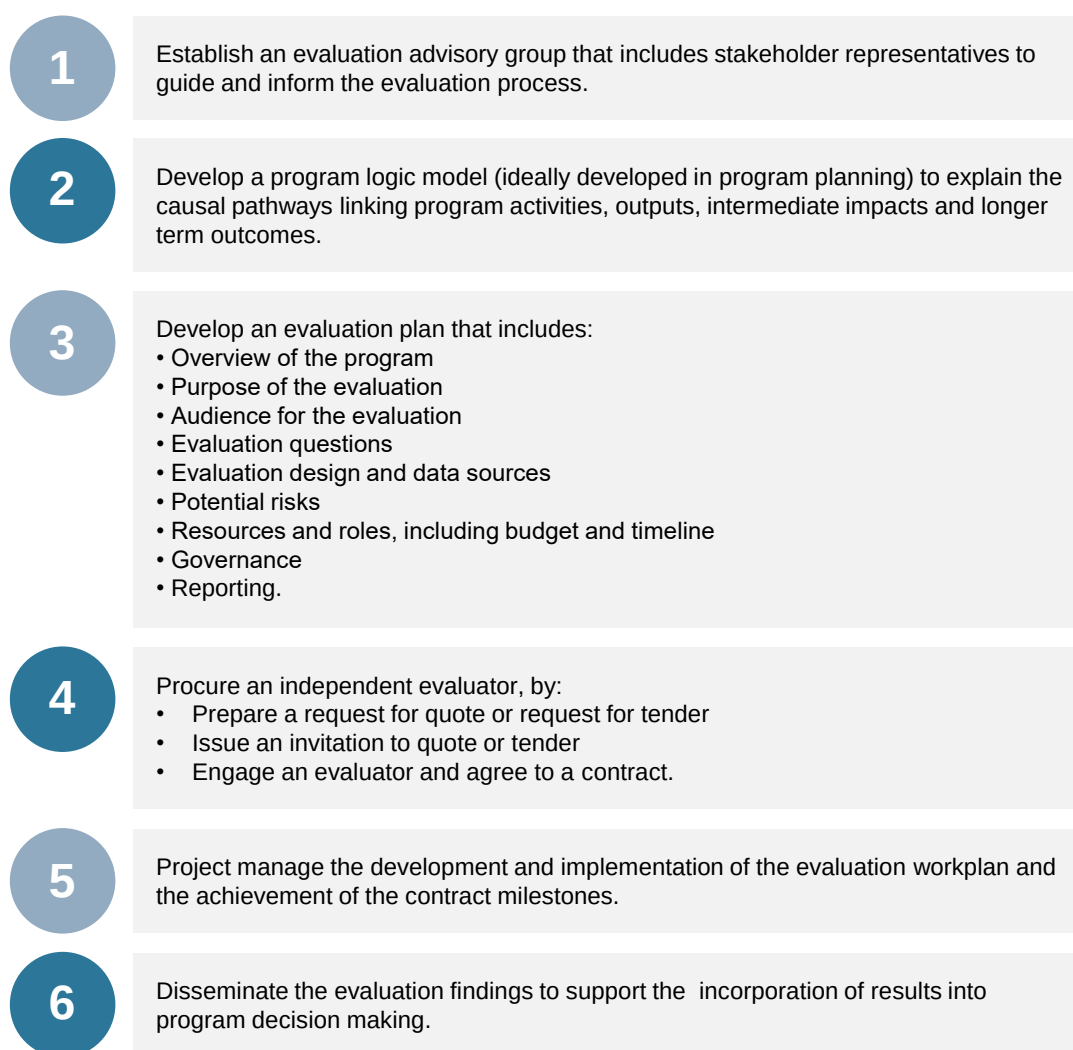
The decision about whether an evaluation or part of an evaluation should be undertaken internally or externally should be based on who is best placed to undertake an evaluation given the evaluation’s circumstances, purpose and approach. However, there should at least be functional independence between evaluators and those who make decisions about policies and program design and delivery.

³⁹ See for example Shelly Bielefeld (sub. 65, p. 2), Eva Cox (sub. 68, p. 3), Accountable Income Management Network (sub. 15, pp. 5-7).

How should evaluations be commissioned?

Once the decision to commission an evaluation has been made, commissioners need to establish arrangements to oversee the evaluation, procure an evaluator and manage the evaluation process (figure 8.2). They may also need to develop a program logic and evaluation plan prior to procuring an evaluator, if these had not been done (as they ideally should have been) as part of the policy design process.

Figure 8.2 **Steps in commissioning an evaluation**



Source: Centre for Epidemiology and Evidence (2019, p. 6).

A request for tender (RFT) or request for quote (RFQ) should outline the specifications of the evaluation project and detail what has been agreed to by policy or program stakeholders. The RFT or RFQ should contain as much information as possible so that evaluators understand the scope of the evaluation and can develop an appropriate response (box 8.6).

Box 8.6 **What to include in a request for tender or request for quote**

A request for tender or request for quote for an evaluation should include:

- an overview of the policy or program including its aims, a description of the policy or program, its policy context, and key stakeholders
- the program logic
- the evaluation plan, including the evaluation's purpose, questions and scope
- any available data sources
- the timeline and an indicative budget for the evaluation
- key deliverables or milestones that need to be met
- the selection process and assessment criteria
- any other requirements as set out in an agency's procurement guidelines or Commonwealth procurement rules.

Source: adapted from Centre for Epidemiology and Evidence (2019).

When procuring an evaluator, agencies may decide to use a panel arrangement, such as the Research, Evaluation and Data Panel or the Collaborative Evaluation, Research and Planning Panel mentioned above. These panels generally include evaluators who have done at least one evaluation for the agency. However, agencies should still ensure that evaluators have appropriate skills, knowledge and experience to deliver high quality results.

When choosing an evaluator, agencies should consider:

- The necessary skills, knowledge and experience needed to undertake the evaluation well. The competencies discussed above may be helpful. It may also be useful to consider an evaluator's performance history.
- Whether the approach proposed by a potential evaluator is feasible, fit for purpose and will deliver a high-quality evaluation. This should take into account the evaluation's purpose and required level of rigour. Agencies may also want to consider the flexibility of the approach, including how innovative or adaptable it may be.
- Value for money, and whether the proposed fee structure and timelines will feasibly deliver a high-quality evaluation. Value for money is not just about price, and agencies must⁴⁰ consider (among other things) quality, fitness for purpose, and the evaluator's experience and work history when assessing value for money.

Once an evaluator has been selected, a contract needs to be agreed. The contract should allow for some flexibility, as evaluation projects usually involve some degree of uncertainty. It should focus on getting the required deliverables to the agreed quality. The contract should also support ethical requirements, including the treatment of intellectual property, ethics clearance and engagement requirements.

⁴⁰ Under Commonwealth procurement rules (Department of Finance 2019, pp. 11–12).

An evaluation is an ongoing process. There should be regular contact between those overseeing the evaluation and evaluators to ensure that the evaluation is being implemented as agreed, any problems are communicated and feedback is incorporated.

When the evaluation report is finalised, the results need to be communicated to key stakeholders, including the participants in the evaluation. This can be done by the evaluators or the commissioners. As outlined in chapter 7, the results of an evaluation should be communicated through publishing a short plain English summary. Results could also be disseminated by:

- running results interpretation or feedback sessions
- sharing results in newsletters or on social media
- developing conference papers or presentations based on the results
- publishing results in peer-reviewed publications.

Commissioners, evaluators, and centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges

Consistent with the overarching principle of the Strategy, it is important that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred in evaluations that are commissioned externally. However, in a commissioned evaluation, responsibility to centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges falls to both evaluators and commissioners. As noted by Kelaher et al. in the context of evaluation ethics:

... the conduct of research does not solely depend on researchers — it is also heavily influenced by the terms of contracts between funders and researchers. Fulfilment of ethical responsibility in Aboriginal and Torres Strait Islander health programs, policies and services therefore needs to incorporate an understanding of the role of all stakeholders, including funding bodies, program implementers and research institutions. (2018b, p. 5)

Given that many decisions, such as timelines and scope, are made before an evaluator comes on board, commissioners need to consider how an evaluation might centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. Questions commissioners should consider are set out in table 8.1.

Table 8.1 Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges — questions for commissioners

<i>Commissioning steps</i>	<i>Questions to consider</i>
Planning for evaluation - Engaging with stakeholders - Establishing evaluation governance - Undertaking evaluation scoping	- How are Aboriginal and Torres Strait Islander people engaged in evaluation planning? - For high priority evaluations, agencies should ensure evaluation governance includes Aboriginal and Torres Strait Islander representation - How do evaluation questions and outcomes reflect those important to Aboriginal and Torres Strait Islander people?
Commissioning the evaluation - Preparing the terms of reference/request for tender (RFT) - Selecting an evaluator	- How should the terms of reference and RFT be shaped by Aboriginal and Torres Strait Islander people — for example, in relation to the required level of engagement, ethics requirements, evaluation outcomes and the treatment of intellectual property? - Does the ethics approval process (where necessary) include Aboriginal and Torres Strait Islander representation? Does the ethics assessment body have expertise appropriate for research and evaluation involving Aboriginal and Torres Strait Islander people? Is the ethics approval process reflected accurately in the RFT? - How are Aboriginal and Torres Strait Islander people engaged in selecting the evaluator? - This may include Aboriginal and Torres Strait Islander representation on the tender assessment panel. - Are Aboriginal and Torres Strait Islander people included as part of the evaluation team (where appropriate)? - Does the selected evaluator have the necessary skills to conduct the evaluation with Aboriginal and Torres Strait Islander people effectively? Do they have existing partnerships with Aboriginal and Torres Strait Islander people?
Undertaking the evaluation - Establishing the evaluation work plan and overseeing its implementation - Monitoring engagement and interaction with the community - Monitoring and maintaining evaluation quality - Providing feedback on the evaluation report	- Is engagement with Aboriginal and Torres Strait Islander people being undertaken as agreed? - Are key ethical milestones (such as securing ethics clearance) being met? - How are participants' contributions of information, knowledges and time being respectfully and appropriately recognised and valued by evaluators and evaluation commissioners? - How are Aboriginal and Torres Strait Islander people engaged in monitoring and assessing the quality of the evaluation? - How are Aboriginal and Torres Strait Islander communities and organisations being supported to provide feedback on the evaluation report?
Disseminating and using evaluation findings - Communicating results - Implementing evaluation recommendations	- How are results communicated back to the Aboriginal and Torres Strait Islander people, communities or organisations engaged in the evaluation? - How are results translated/interpreted to be culturally appropriate? - For high priority evaluations, agencies may want to require results to be translated either by an Aboriginal or Torres Strait Islander evaluator or by a non-Indigenous evaluator in partnership with Aboriginal and Torres Strait Islander people/communities/organisations. - How are management responses communicated back to the community and organisations?

Agencies also need to consider how their commissioning process centres Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. An example might be 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. Kelaher et al. (2018b) identified examples of good practice when commissioning evaluations (in the context of health programs), a selection of which are presented in box 8.7. The examples point to the importance of:

- including an evaluator's experience in conducting research and their relationships with the communities involved as part of the selection criteria
- commissioners providing sufficient funding and time to allow for appropriate engagement
- commissioners and evaluators working together to ensure ongoing communication with communities, and demonstrating that communities' contributions have been heard
- ensuring that contracts and agency approval processes do not prevent findings from being communicated back to the community.

For mainstream policies and programs, agencies should consider an evaluation's purpose and potential impact on Aboriginal and Torres Strait Islander people when thinking through how to address the questions raised in table 8.1. Box 8.8 provides a hypothetical example of one approach an agency might take when commissioning an evaluation of a mainstream program that affects Aboriginal and Torres Strait Islander people. However, it is important to note that there are many ways mainstream evaluations can be approached.

As discussed in chapter 7, agencies should report on how an evaluation aligns with the principles of the Strategy. Where an evaluation (or part of an evaluation) has been commissioned, this reporting should extend to how the principles were upheld in any commissioning arrangements. This means that a degree of accountability will be placed on agencies to ensure they are upholding the principles of the Strategy in their commissioning arrangements. It will also provide information that can inform the monitoring of agencies' performance in implementing the Strategy.

Box 8.7 **Examples of how Aboriginal and Torres Strait Islander people are centred in a commissioned evaluation**

Two gathering places in the eastern metropolitan region of Melbourne

Gathering places were local community hubs, which were community controlled and led. They aimed to provide a cultural place for Aboriginal communities to connect and heal. An evaluation was commissioned by the Victorian Department of Health and undertaken by Onemda.

Positive practice

- The Victorian Department of Health had a strong focus on ensuring evaluators had the right skillset to undertake the evaluation. They had three criteria — the proposal is both technically and culturally appropriate; key personnel undertaking the project can demonstrate research and evaluation experience with Victorian communities; and the bidder had recent work experience and existing relationships with the communities engaged in the evaluation.
- The evaluation team comprised of three researchers, two of whom were Aboriginal. They established a reference group with local leaders.
- The evaluation used participatory opportunities to engage the community in analysing, making sense of and validating the findings of the evaluation.

Limitations/challenges

- The engagement strategy was not resourced and there was an expectation that the community would participate despite the burden on them.
- Departmental embargoes on publishing the evaluation meant that results were not able to be shared with the community until a year after the evaluation was submitted to the Department.

Stronger Communities for Children program

Stronger Communities for Children was a community development program run by the Department of the Prime Minister and Cabinet. It aimed to build stronger and safer communities in remote areas in the Northern Territory by pairing communities with a facilitating partner to deliver services to support local children and families. It was evaluated by Ipsos and Winangali.

Positive practice

- Tender criteria required evaluators to have experience working in remote Indigenous communities.
- An Aboriginal company was contracted to provide cultural advice and to lead engagement with the communities.
- The evaluation built on already on-going partnerships and evidence-generation activities as part of the program's implementation.
- Participants were provided with detailed evaluation reports as well as two-page plain language summaries. A knowledge sharing seminar was also undertaken.

Limitations/challenges

- The evaluation faced challenges in obtaining ethics clearance, as the existing ethics process was set up to review research and not evaluations. The ethics process was unsure of how to address the developmental approach and conversational interviewing methodology.

Source: Kelaher et al. (2018b).

Box 8.8 A hypothetical commissioned evaluation of a hypothetical business support program

A hypothetical business coaching program is aimed at helping small to medium business enterprises expand. It is not targeted at Aboriginal and Torres Strait Islander businesses. The evaluation is not a high priority evaluation under the Indigenous Evaluation Strategy, but program managers still want to investigate how to improve the program for Aboriginal and Torres Strait Islander and other users.

- Aboriginal and Torres Strait Islander people were asked to provide feedback on the evaluation plan. Their feedback noted that access to the program might be a problem. They also noted that some businesses might choose to employ more people rather than growing profits, and so the value of the program should not be assessed based on its impact on profits alone.
- 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. Feedback suggested location A. The commissioners approached several businesses in location A ahead of time and they agreed to participate. This was noted in the request for tender.
- The commissioner decided to divide up the evaluation commissioning, with one evaluator contracted specifically to do the case study and another evaluator contracted to undertake the general evaluation as well as the data analysis. The case study evaluator demonstrated that they had experience working with Aboriginal and Torres Strait Islander people, having undertaken similar case studies in remote Australia.
- 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 performed as well as non-Indigenous businesses, smaller businesses were less likely to apply for the grant. The case study then explored these issues in detail with two different businesses.
- It was written into the contract 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 business was given a short summary of the findings and their data.

9 Data for evaluation

Key points

- Good data are essential for high quality evaluation and need to be planned for early (at the policy or program design stage).
- Evaluations can use existing data (administrative, monitoring and survey data) and/or new data collected as part of the evaluation (such as surveys, interviews or focus groups). A review of recent Australian Government department evaluations showed that most collected new data, rather than using monitoring or other existing data.
- A key concern raised about data collected for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people was the lack of engagement with Aboriginal and Torres Strait Islander people in decisions about data collection, use and governance (and what this means for the relevance of, and community trust in, the data). Other concerns were about data burdens (on participants and service providers) and data quality (including that monitoring data are often focused on outputs rather than outcomes, and that baseline data and comparison groups are commonly missing in evaluations).
- Planning early for an evaluation allows for data requirements to be built into policy and program design. This can improve the quality of evaluation by allowing the collection of baselines and counterfactuals, and consideration can be given to appropriate sampling strategies before a policy is in place. Evaluation planning should consider the:
 - data needed to answer evaluation questions
 - data needed to produce credible results (baselines, counterfactuals, sampling methods, the use of quantitative and/or qualitative data)
 - availability and suitability of existing data for the evaluation, and what additional data should be collected.
- Other measures that will improve data for evaluation include:
 - engaging with Aboriginal and Torres Strait Islander people on data planning and management
 - developing an evaluation data dictionary with core indicators that Aboriginal and Torres Strait Islander people value, and providing guidance to agencies on consistent data collection practices
 - making better use of existing data.

Good data are essential for high quality evaluation. As Deborah Cobb-Clark said:

... evaluations always rest on the available data. It is simply not possible to evaluate what we cannot observe. It is not uncommon for data limitations to constrain the evaluation questions, the evaluation method, the quality of the evaluation, and indeed whether an evaluation is even

possible. ... One of the most important investments we can make is in data sources which can be used to support public policy evaluation. (2013, p. 84)

And early planning for what data should be collected for evaluation — at the policy or program design phase — is critical for good data. If left until after a policy is in place, it can limit the conduct of an evaluation and can make the collection of data more expensive. As the Magenta Book said:

Without appropriate planning of data collection or data access an evaluation may be impossible, severely limited, or unnecessarily expensive. If the data collection is poorly designed, this may result in inaccurate data being collected and false inferences being drawn from the evaluation. (HM Treasury 2020c, p. 53)

This chapter looks at:

- how data are currently used in Australian Government agencies' evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people (section 9.1)
- issues raised by participants about data for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people (section 9.2)
- ways to improve data for evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people (section 9.3).

9.1 How are data currently used in evaluations?

Evaluations can use existing data — administrative, monitoring and survey data — and new data collected as part of the evaluation (such as surveys, interviews or focus groups) (box 9.1).

Decisions about what data to collect for an evaluation should be based on what data are required to answer the evaluation questions (with input from those affected by the policy or program and evaluation users) and what data are already collected and available. While existing administrative and other data are cost-effective options (and using existing data can mean a lower burden on program participants), new data are also often required to answer specific evaluation questions.

Box 9.1 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 (chapter 4 looks at the strengths and weaknesses of the different methods).

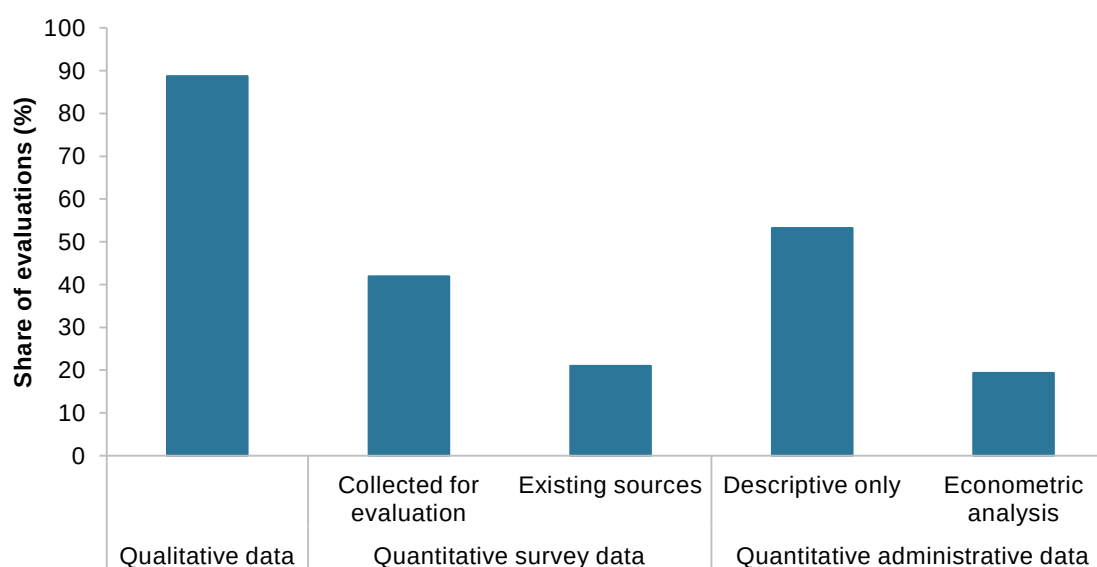
Source: HM Treasury (2020c).

Evaluations rely heavily on new data

As discussed in chapter 3, most of the evaluations provided through the Commission's information request (appendix B) used both qualitative and quantitative data. The Commission also looked in depth at the five most recent evaluations published by Australian Government departments and the National Indigenous Australians Agency (NIAA) and found that most of these evaluations collected new data rather than using existing data from surveys or administrative monitoring data (figure 9.1).

Figure 9.1 Data sources used in recent Australian Government department evaluations^a

Per cent of evaluations that report using each data type^b



^a Based on a review of 57 department evaluations (the 5 latest evaluations undertaken by the department) and 5 National Indigenous Australian Agency (NIAA) evaluations (NIAA is the only agency included in the sample). Where departments had not completed the information request (chapter 3), evaluations were taken from departments' websites. Only published evaluations were used. The sample includes evaluations of both Indigenous-specific policies and programs and 'mainstream' policies and programs (including both those that examined impacts on Aboriginal and Torres Strait Islander people and those that did not). ^b Most evaluations report using multiple data types.

Source: Productivity Commission analysis.

Quantitative data used in the evaluations were from a range of sources, with a greater reliance on existing data sources than for qualitative data, which were almost all collected specifically for evaluation. However, relatively few evaluations used existing data sources to conduct econometric or inferential analysis.

- About 40 per cent of evaluations used quantitative survey data collected during the evaluation process. In most cases, these were surveys of program participants, but in some evaluations service providers and other stakeholders were surveyed.
- A small number of evaluations (about one-fifth) used data from existing quantitative surveys. These included regular surveys of providers and jobseekers undertaken by the Department of Employment, as well as Australian Bureau of Statistics (ABS) survey data. However, it is worth noting that evaluations made relatively little use of ABS survey data beyond providing contextual information, such as describing the characteristics of regions being examined in the evaluation.
- Existing sources of quantitative administrative and monitoring data were used in about 70 per cent of evaluations, but typically only to provide descriptive statistics, such as describing the number or demographic characteristics of participants in a program, or aggregate characteristics of a region.

-
- Only about 20 per cent of evaluations used existing administrative or monitoring data for econometric analysis or inferential statistics.

9.2 Some issues raised about data for evaluation

Data collection and analysis can be costly and burdensome

There are costs to collecting (including costs to program providers and participants) and using data in evaluation. A number of Aboriginal and Torres Strait Islander organisations spoke about the burden that data collection and analysis for evaluations imposes on service providers and participants (box 9.2).

Box 9.2 Some concerns raised about the burden of collecting and reporting data

National Aboriginal and Torres Strait Islander Legal Services:

Funding often sees ACCOs [Aboriginal Community Controlled Organisations] burdened with reporting requirements, short-term funding and unsustainable 'pilot' funding. Evaluation overburden may also include requiring organisations to report to separate agencies on programs with overlapping objectives, while the data and learnings from evaluation may not be shared. (sub. 97, p. 13)

Victorian Aboriginal Child Care Agency:

Current practice is that frontline Aboriginal services and Aboriginal clients bear a considerable burden in relation to evaluation activities. They are often responsible for engagement of participants, identification and management of ethical risks, data collection and utilisation of evaluation findings, without necessarily having been involved in decisions to determine evaluation priorities and evaluation design. (sub. 26, p. 2)

First Nations Media:

A collaborative approach to data collection should be taken to minimize the additional work expected of Aboriginal and Torres Strait Islander people who are disproportionately burdened with the continual provision of information about multiple aspects of their lives. (sub. 30, p. 13)

A recent review of Indigenous-specific Primary Health Network (PHN) datasets also found that reporting requirements outside of the PHN datasets add to the burden on service providers. The review noted that some organisations are required to complete 50 reports a year, and concluded that better coordination and negotiation within and across departments, as well as inclusive data governance, are needed to reduce reporting burdens (AIHW 2020).

The burden of collecting data can be exacerbated for providers when data provided are not used by government, are not useful to the data providers themselves, or do not contribute to improved outcomes. Queensland Aboriginal and Islander Health Council, for example, said:

More often than not, the information provided by ATSICCHOs [Aboriginal and Torres Strait Islander Community Controlled Health Organisations] is rarely translated into improved support or services for their benefit. (sub. 55, p. 7)

Questions were also raised about whether the right data are collected

Data may not reflect what Aboriginal and Torres Strait Islander people value

The Commission heard that the data collected for monitoring and evaluation do not always include metrics and indicators that are relevant to, or of high priority for, the lives of Aboriginal and Torres Strait Islander people (box 9.3).

Box 9.3 Concerns that the data collected do not reflect what is important to Aboriginal and Torres Strait Islander people

Institute for Urban Indigenous Health:

In terms of current evaluation approaches and methods to evaluate Indigenous programs, the research suggests a current dearth of culturally responsive outcome measures for use with Aboriginal and Torres Strait Islander peoples. Most outcome measurement tools have been developed with participants from Western backgrounds. The appropriate evaluation of health services for Aboriginal and Torres Strait Islander peoples demands culturally responsive goal-setting tools that can be used by inter-professional teams. (sub. 53, p. 21)

National Aboriginal and Torres Strait Islander Legal Services:

... current models of evaluation do not necessarily capture elements such as 'connection to land, community and culture' as part of successful outcomes of programs. (sub. 97, p. 9)

Patrick Sullivan et al:

If the overall goal is to improve outcomes for Indigenous people, then they need to have a say in whether, from their perspective, outcomes have improved. ... our research has found that program goals may be met, but outcomes have not improved, particularly (but not only) from the perspective of those whose lives are meant to be improving. (sub. 11, p. 2)

Tony Dreise et al:

... while we endorse the use of well-designed quantitative research approaches where appropriate, we stress that a broader range of variables be considered in order to capture Indigenous rights, perspectives and aspirations within evaluation. (sub. 33, p. 8)

Australian Council of TESOL Associations:

There are large gaps in data on the use of Indigenous languages in Australia for reasons of under-counting, over-counting, and the fact that some languages do not have widely used names. (sub. 87, p. 46)

Where this is the case, observed improvements in indicators may not necessarily reflect improvements to Aboriginal and Torres Strait Islander people's lives. Adverse or unintended consequences of policies and programs affecting Aboriginal and Torres Strait Islander

people could also go unnoticed if relevant data are not collected. Fiona Allison and LawRight highlighted an evaluation of a health justice project that added a range of variables of interest to Aboriginal and Torres Strait Islander people to those collected as part of standard monitoring data for the project:

The client database we have drawn quantitative data from is not Indigenous specific. It is used across all CLCs [Community Legal Centres]. ... As an example, outcomes to be measured in the evaluation now include indicators such as strengthening of cultural determinants of health (strengthening of cultural and community connection); feeling heard and validated in a culturally capable way; and addressing racism and social exclusion based on race. These (positive) justice and access to justice outcomes have been identified as important measures within an Indigenous-focused HJP [health justice project] evaluation context. (sub. 18, p. 4)

Concerns were also raised that outcome measures used in evaluation often are not strengths-based, are insufficiently contextualised, or focus unnecessarily on comparisons between Aboriginal and Torres Strait Islander people and non-Indigenous people (Fogarty et al. 2018; Walter 2016). The Tasmanian Government said:

[Framing evaluation questions around strengths] is necessary to avoid reporting solely on negative outcomes without context and an explanation of the social determinants of negative outcomes, including the effects of inter-generational trauma, discrimination and racism. (sub. 100, p. 2)

A number of recent Indigenous-led studies have sought to improve the collection of data relevant to Aboriginal and Torres Strait Islander people's lives. These include the *Mayi Kuwayu* study (box 9.4) and the *Community Wellbeing from the Ground Up: A Yawuru example* research project (Yap and Yu 2016a). While these projects are tailored to the needs of the communities in which they are undertaken, the data collected cover topics including cultural connectedness and identity, family relationships, experiences of racism, cultural safety and voice.

Box 9.4 **The Mayi Kuwayu study**

The Mayi Kuwayu study is a national study of the cultural determinants of health of Aboriginal and Torres Strait Islander people, funded by the Lowitja Institute and the National Health and Medical Research Council. The study arose from community-identified priorities and developed measures of culture and identity with Aboriginal and Torres Strait Islander communities across Australia. It aims to be longitudinal with follow-up surveys being conducted every 3-5 years and will link to survey participants' Medicare records. In addition to collecting data on demographics and health/wellbeing outcomes and behaviours, the survey collects data on:

- cultural practice and expression — connection to Country, beliefs and knowledge, cultural expression, self-determination and leadership, language, family, kinship and community, and identity
- experiences and environments — services in the community, experiences of racism, community safety, environmental conditions, and life events
- family support and connection — family cohesion and connectedness, caring for others, the Stolen Generation.

An Aboriginal and Torres Strait Islander committee provides oversight and ensures appropriate Indigenous governance. A data governance committee will also approve the use of data once collection is finalised. There is ongoing engagement through newsletters, forums and social media communication.

Sources: Jones et al. (2018); Mayi Kuwayu (2019).

Monitoring data are often focused on outputs or compliance rather than outcomes

Data for evaluation should aim to measure how programs contribute to improved outcomes. However, many evaluations use data that focus on outputs rather than outcomes (chapter 3). This makes it difficult to assess what is working for whom, and in what circumstances. The National Justice Project also suggested that:

A focus on outputs or 'targets' rather than impacts and outcomes also creates a dangerous culture or discourse that 'blames' the people meant to benefit from a policy or program, rather than the design and implementation of the policy or program itself. (sub. 51, p. 4)

A focus on outputs rather than outcomes can come about because:

- data collection and evaluation are not planned during policy and program design and are not built into policy or program implementation
- outcomes can be difficult to measure, especially where outcomes are intangible or are affected by multiple policies and programs
- agencies adopt an overly risk-averse, compliance-focused approach to data collection
- input, output and outcome data are collected by different agencies and there are challenges in obtaining approval to share or link data
- improvements to outcomes can only be seen in the long term (PC 2017b, chapter 8).

Lack of transparent data on inputs

Data on policy and program inputs — what policies and programs operate where, and how much funding they receive — are important for undertaking rigorous evaluations and can:

- provide important context to an evaluation about which programs operate in an area
- assist in assessing causality when there are multiple programs operating in an area
- identify the funding in a community to ascertain duplication and waste or returns on investment.

Some participants pointed out the lack of information on inputs to programs.

A number of the issues that we have encountered in relation to baseline data are well known. These include the poor quality of remote area data, lack of economic development data for Indigenous people, incomplete or undeveloped regional services or program investment data, absence of data relating to culture and language, lack of concordance of geospatial data (and there are many more). (Empowered Communities, sub. 41, p. 8)

Without baseline data which gives us a picture of the current service system the Indigenous Evaluation Strategy runs the risk of being limited by program, jurisdiction or Department. A baseline data set would allow for thematic, programmatic or individualized service evaluation. (Generation One, Minderoo Foundation, sub. 5, p. 4)

While individual agencies have information about the policies and programs they deliver, they may lack data on policies and programs delivered by other Australian Government agencies, state, territory and local governments or non-governmental organisation. This lack of data and transparency about policy and program inputs can also hinder community-led evaluation efforts.

The Commission produces, for the Steering Committee for the Review of Government Service Provision, high-level estimates of expenditure on Aboriginal and Torres Strait Islander people. These data have limited usefulness for evaluation for several reasons:

- reporting does not link expenditure to outcomes and, as such, cannot be used to determine the efficiency of spending
- data are not provided at a sufficiently detailed level of geography to inform evaluations
- the categorisations of expenditure are too broad to usefully inform evaluations, and data are not available at the policy or program level.

The Commission's study into *Expenditure on Children in the Northern Territory* tried to track expenditure on programs and encountered several problems, including the use of different service definitions and geographic standards, many of which were not granular enough to map services or expenditure to a particular community (PC 2020, chapter 3).

Baseline and comparison group data are not always collected

Comparing outcomes before and after a program is implemented, or between program participants and non-participants with similar characteristics, is important for rigorous impact evaluation. However, often baseline and comparison group data are not collected prior to a policy or program commencing, making it difficult to understand how outcomes have changed and to make causal inference about whether the policy or program affected these outcomes. As the Western Australian Government said:

The absence of baselines, limited data or inability to access datasets are common reasons why evaluations are not completed in a timely manner or key evaluation questions cannot be answered. (sub. 74, p. 5)

The Commission examined several evaluations that identified a lack of baseline or comparison group as a significant limitation (AHA 2018; KPMG 2019; Mitchell, Bandara and Smith 2018). Often a lack of pre-program data was the cause of the problem. For example, trial data from the Department of Social Services' Individual Support Placement trial only covered the final two years of the three-year trial.

In some cases, it can be difficult to collect data on baselines and comparison groups where programs are small or a clear counterfactual cannot be identified. For example, in the evaluation of the former Department of Jobs and Small Business' Relocation Assistance Program a comparison group could not be identified as there were differences between those that decided to move and those that did not that could reasonably have influenced outcomes (DJSB 2018).

Collected data may not be of sufficient quality

Monitoring data can be poor quality

Ideally, data collected as part of policy and program implementation would be used in evaluation, saving time and resources that would otherwise be devoted to collecting evaluation-specific data. However, when reviewing Australian Government evaluations (section 9.1), the Commission found several examples where monitoring data were of poor quality and were unable to be used in the evaluation (box 9.5). Some of the issues included:

- inconsistent outcome measures and data collection practices, which led to a lack of comparability across jurisdictions, service providers, or key administrative datasets
- missing data on key outcome and demographic variables
- delayed data availability, with evaluators unable to access monitoring data in time to complete evaluations
- program data not covering the entire program length
- inaccuracies in data entry or data transfer procedures

- changes over time in variables, definitions, software or systems which made long-term comparisons difficult.

Box 9.5 Limitations of monitoring and administrative data identified in recent Australian Government evaluations

***An Evaluation of the Family Advocacy and Support Services* (Inside Policy 2018)**

Some key variables were not available in administrative data. Differences in data collection methods made comparisons across jurisdictions difficult.

***Tackling Indigenous Smoking program: Final Evaluation Report* (Mitchell, Bandara and Smith 2018)**

Differences in outcome measures and data collection approaches meant that monitoring data could not be aggregated across jurisdictions. There was a lack of standardised outcome measures, a lack of baseline data and no comparison group.

***Evaluation of the Partners in Culturally Appropriate Care Program* (AHA 2018)**

Activity report data collected by service providers differed in the level of detail provided. Not all activities undertaken by providers were listed in activity reports. There were no pre-program data available and reports on changes over time were based on recall.

***Evaluation of the Pilot Program of Specialist Domestic Violence Units and Health Justice Partnerships Established under the Women's Safety Package* (Social Compass 2019)**

There were inconsistencies in the way service providers collected and reported on data in quarterly reports.

***Settlement Grants Program* (Smyth et al. 2017)**

Key outcome and demographic data were collected on a voluntary basis so there were many missing and 'not stated' values. Different classification systems for regions, language and country of birth were used across different administrative data sources, making comparisons difficult.

***NDIS [National Disability Insurance Scheme] Evaluation Consolidated Report* (Mavromaras et al. 2018)**

The evaluation had originally planned to use administrative data linked with survey data, but evaluators were not able to access the administrative data during the course of the evaluation.

***Final Report for the Evaluation of the Individual Placement and Support Trial* (KPMG 2019)**

Administrative data were not available for the whole length of the trial. There were also inaccuracies in data entry, discrepancies between different administrative data sources and missing data on Indigenous status and other key variables.

***Job Commitment Bonus for Young Australians Evaluation Report* (DJSB 2019)**

The introduction of a new reporting system in the middle of the evaluation period resulted in data discrepancies. Reported data on exits from income support differed across administrative datasets.

***Evaluation of Job Services Australia 2012–2015* (DESSFB 2020)**

Some outcome measures changed when a new employment services model was introduced, reducing data comparability over time.

Sample sizes can be too small or not representative

Aboriginal and Torres Strait Islander people are typically a small proportion of participants in ‘mainstream’ policies and programs, which means sample sizes can be an issue for evaluators (Muir and Dean 2017). Without sufficient planning, a random sample of participants may not yield enough Aboriginal and Torres Strait Islander respondents to produce meaningful results. There can also be barriers for participants to complete surveys or be involved in interviews, including language, Internet access, illiteracy and high geographic mobility. These barriers can contribute to lower response rates and higher attrition rates that further reduce sample sizes.

Even where sample sizes are sufficiently large to analyse impacts on Aboriginal and Torres Strait Islander people, using aggregated data can mask the diversity of their experiences. As the Smith Family noted:

The [Indigenous Evaluation Strategy] will also need to acknowledge the diversity of characteristics, experiences and perspectives of Aboriginal and Torres Strait Islander peoples and acknowledge that aggregate data may mask significantly different experiences, by for example, socioeconomic background, location and gender. (sub. 56, p. 15)

Sample sizes are not a problem unique to evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. The Commission identified several evaluations that struggled with low response rates for the general population, particularly when trying to collect longitudinal data from participants or service providers (KPMG 2019; Mavromaras et al. 2018; ORIMA Research 2017). Poor survey design and difficulties in gaining access to program participants or service providers also affected the representativeness of samples (AHA 2018; Social Compass 2019).

Sample sizes are also not a problem unique to data collected for evaluations. Some existing data sources have small Aboriginal and Torres Strait Islander sample sizes or experience attrition issues. For example, the Household, Income and Labour Dynamics in Australia (HILDA) survey has an insufficiently large sample size to usefully examine impacts for Aboriginal and Torres Strait Islander people. It also has high rates of attrition for Aboriginal and Torres Strait Islander people (Biddle and AIFS 2014; Watson and Wooden 2004).

However, appropriate design — including deliberate oversampling — and engagement with Aboriginal and Torres Strait Islander people, can improve responses and increase sample sizes. For example, many academics point to the success of the Longitudinal Study of Indigenous Children (LSIC). As noted by Thurber et al.:

... the employment of Indigenous interviewers, dedicated community consultation process and continuous feedback loop are likely to underlie the ability of the LSIC [Longitudinal Study of Indigenous Children] study to maintain data integrity and minimize attrition while ensuring communities’ good will towards the study. (2015, p. 797)

Data on Indigenous status are not consistently collected

Identifying the impacts of policies and programs on Aboriginal and Torres Strait Islander people relies on being able to accurately identify Aboriginal and Torres Strait Islander program participants.

Data on Aboriginal and Torres Strait Islander status are not collected in all government administrative data, preventing their use for evaluating impacts on Aboriginal and Torres Strait Islander people. For example, in an evaluation of the National Human Papillomavirus Program commissioned by the Department of Health, evaluators were unable to examine vaccination rates for Aboriginal and Torres Strait Islander people in the Northern Territory or Queensland because Indigenous status was a non-mandatory variable in the National Human Papillomavirus Vaccination Program Register (NCIRS 2014).

Where data on Aboriginal and Torres Strait Islander status are collected, most data sources rely on respondents to self-identify as Aboriginal and/or Torres Strait Islander. For Aboriginal and Torres Strait Islander people, the decision whether to identify their Indigeneity for data collection purposes is — rightly — theirs to make (box 9.6), but it presents challenges for researchers and evaluators when an individual's Indigenous status varies within and across datasets, particularly when tracking outcomes over time.

There have been attempts to standardise the collection of statistics on Indigenous status. The ABS developed a standard question on Indigenous status for the 1996 Census:

Question: Are you of Aboriginal and Torres Strait Islander origin?

Response categories: Yes – Aboriginal; Yes – Torres Strait Islander; No [note: respondents can choose to identify as being of both Aboriginal *and* Torres Strait Islander origin]. (ABS 2014)

The Australian Institute of Health and Welfare (AIHW) has also developed best practice guidelines outlining how the standard question should be asked and recorded, and how to put the guidelines into practice (for example, by providing training to staff and improving cultural capability) (AIHW 2010).

As part of the Closing the Gap agreement, the Australian, state and territory governments all agreed to adopt the standard question and provide training to data collectors in education and health (COAG 2009, Schedule F). As such, the ABS question has now become the standard for surveys and most administrative data collections.

However, some government programs — including ABSTUDY and the Indigenous Procurement Program — use the legal definition of Indigeneity, which requires that individuals have Aboriginal or Torres Strait Islander descent, identify as Aboriginal or Torres Strait Islander, and are accepted by the community as being Aboriginal or Torres Strait Islander (DPMC 2019f; Services Australia 2019). The Indigenous Procurement Policy requires that 50 per cent of a business' owners meet the legal definition of Indigenous.

Box 9.6 Indigenous identification in data

Many data sources rely on respondents self-identifying if they are Aboriginal and/or Torres Strait Islander. The decision whether to identify as Indigenous in response to a survey or administrative data request is often complex. As the Australian Bureau of Statistics (ABS) said:

The decision to disclose one's Indigenous status is a personal one, and potentially complex. In addition to an individual's assessment of the question and the data collection context, identification may be influenced by attitudes, beliefs and behaviours that the individual is not consciously aware of. (ABS 2013, p. 6)

Many factors influence a person's choice to self-identify. The ABS (2013) and Aboriginal Affairs (NSW) (2015) have investigated factors that affect a person's propensity to identify. Both studies cited similar reasons to identify. Positive influences included:

- pride in their Aboriginal and Torres Strait Islander heritage
- an understanding of the benefits of identifying, such as improved funding allocation, better services, and greater recognition of Aboriginal and Torres Strait Islander-specific issues
- having a 'Confirmation of Aboriginality'^a (which can help people feel more confident to identify).

Negative influences included:

- past negative experiences in identifying, such as experiences from the Stolen Generation or casual racism in service delivery
- a belief that identifying may have negative repercussions for the community — that these statistics would be communicated in a negative way, perpetuating racism and discrimination
- a lack of information about why the data are being collected.

Differential or changing propensity to identify as Aboriginal and Torres Strait Islander can have a pronounced impact on statistics. For example, between the 2011 and 2016 Censuses, the population of Aboriginal and Torres Strait Islander people rose by 18.4 per cent, with about one-fifth of this change unable to be explained by demographic factors such as births, deaths and migration. A changing propensity to self-identify as Aboriginal and/or Torres Strait Islander has been acknowledged as one possible factor that is contributing to the increase in the Aboriginal and Torres Strait Islander population recorded in the census (ABS 2018).

^a This is a document confirming an individual's Aboriginal or Torres Strait Islander heritage.

As well as differences in the questions used to collect data on Indigenous status, there is variation in how data are collected across states and territories, as well as between datasets. These include differences in who asks and answers questions about Indigenous status, whether answering Indigenous status questions is voluntary or mandatory, and whether respondents can change their Indigenous status in longitudinal datasets.

Data collection is not always planned for

Many of the data problems identified stem from poor planning for data collection during policy and program design and implementation. As the Western Australian Government said:

... data collection is frequently an after-thought in policy and program design ... clearly articulating data requirements and processes at the commencement of the project or policy can help to mitigate this difficulty. (sub. 74, p. 9)

One example is the Cashless Debit Card evaluation commissioned by the Department of Social Services in 2016 (box 5.2). This evaluation was commissioned only a few weeks before the program was implemented, which meant that there was not enough time to collect adequate baseline data. The Australian National Audit Office (ANAO) said:

The commissioning of ORIMA on 22 February 2016 (22 days before the trial commenced) allowed little time to develop, test, agree and establish baseline data to measure against the KPIs [key performance indicators]. The KPIs were developed by ORIMA, in consultation with stakeholders, over the first few months of the trial and provided to Social Services on 26 July 2016, over four months after the trial had commenced. (2018, p. 37)

Concerns about engagement with Aboriginal and Torres Strait Islander people in data planning and governance

Several of the problems with data collection and use identified above have their origins in a lack of engagement with Aboriginal and Torres Strait Islander people in data planning, collection and governance. Many participants pointed to such a deficiency in current data practices. Maggie Walter, for example, said:

Aboriginal and Torres Strait Islander people are treated as objects of data, and we have no say in what is collected, where and how it is collected, who uses it, and for what purposes, under what circumstances and what comes from it. (sub. 112, p. 4)

The NIAA pointed to the benefits of data governance arrangements that actively engage with Aboriginal and Torres Strait Islander people:

Good governance of Indigenous data enables the collection and availability of data relevant to the priorities of Indigenous peoples and communities. A crucial component of good data governance is the active involvement of Aboriginal and Torres Strait Islander people in making decisions concerning the management of Indigenous data. (sub. 86, p. 6)

As well as its impact on the relevance of data collected, the lack of engagement with Aboriginal and Torres Strait Islander people on data collection, management and governance is likely to have implications for data quality. For example, individuals may choose to not identify as Aboriginal or Torres Strait Islander, answer questions incorrectly, or not complete questions at all for fear that their answers may be used in a negative way (Aboriginal Affairs (NSW) and Department of Education (NSW) 2015).

The calls for greater Aboriginal and Torres Strait Islander engagement in data governance are part of a wider push, both within Australia and overseas, to improve governance of data relating to Indigenous peoples and move towards Indigenous data sovereignty (box 9.7).

Box 9.7 What is Indigenous data sovereignty?

Indigenous data sovereignty is defined as the right of Indigenous peoples to exercise ownership over Indigenous data, where Indigenous data are defined as information or knowledge, in any format or medium, which is about and may affect Indigenous peoples both collectively and individually. Ownership of data can be expressed through the creation, collection, access, analysis, interpretation, management, dissemination and reuse of Indigenous data.

At an Australian Indigenous data sovereignty summit held in 2018, delegates asserted that Aboriginal and Torres Strait Islander people have the right to:

- exercise control of the data ecosystem including creation, development, stewardship, analysis, dissemination and infrastructure
- data that are contextual and disaggregated (available and accessible at individual, community and First Nations levels)
- data that are relevant and empower sustainable self-determination and effective self-governance
- data structures that are accountable to Indigenous peoples and First Nations
- data that are protective of, and respect, the individual and collective interests of Aboriginal and Torres Strait Islander people.

The Global Indigenous Data Alliance has also developed the CARE principles for Indigenous data governance.

- **Collective benefit** — data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.
- **Authority to control** — Indigenous peoples' rights and interests in Indigenous data must be recognised, and their authority to control such data must be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, as well as Indigenous lands, territories, resources, knowledges and geographical indicators, are represented and identified within data.
- **Responsibility** — those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples' self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples.
- **Ethics** — Indigenous peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.

These principles are intended to sit alongside the FAIR (Findable, Accessible, Interoperable, Reusable) principles for open data.

Sources: Maïam nayri Wingara and AIGI (2018); Research Data Alliance International Indigenous Data Sovereignty Group (2019).

9.3 Improving data for evaluation

There are a number of ways data collection, quality and governance could be improved (figure 9.2). These include:

- planning for evaluations early and ensuring data collection processes are in place
- engaging with Aboriginal and Torres Strait Islander people on data planning and management in evaluations, as well as more broadly
- developing an evaluation data dictionary with core indicators of what Aboriginal and Torres Strait Islander people value, and providing guidance on consistent data collection practices
- improving the use of existing data in evaluation.

Figure 9.2 Improving data for evaluation — problems and solutions

	Planning early for data collection	Engaging with Aboriginal and Torres Strait Islander people on data	Developing a data dictionary	Better use of existing data
Data collection is costly and burdensome				
The right data are not collected				
Data are of poor quality				
Lack of engagement with Aboriginal and Torres Strait Islander people				

Planning early for evaluation and data collection

The benefits of early planning

There are many advantages of identifying data needs for evaluation prior to a policy or program commencing. Appropriate baseline and counterfactual data can be collected. Data can be collected over time as part of the implementation of policies and programs. Approvals to share and integrate data (including consent from participants) can be obtained ahead of the evaluation. As the 2011 Magenta Book stated:

[Retrospective data collection] is very likely to be more expensive than collecting data at the same time the policy was taking place. In addition, data may no longer be available or may be inaccurate or piecemeal and the opportunity to validate this data may have been lost. Information may not have been collected on drop-outs which may bias the findings. This is particularly relevant where this information is required to contact participants or where it is needed in order to identify them in other datasets. In summary, it can mean that an evaluation is not possible or that its findings are much less reliable than if data had been collected at the same time the policy was being delivered. (HM Treasury 2011, p. 77)

What to consider during data planning

Data planning should be part of evaluation planning processes (discussed in chapters 5 and 7) as data needs will be informed by the evaluation questions, scope and approach. Data planning should consider:

- what data are needed to answer each evaluation question?
- what data are needed to produce credible results (including the use of quantitative and/or qualitative data, and sampling methods)?
- what existing data are available and suitable for the evaluation and what additional data should be collected?

The Department of Industry, Innovation and Science's Evaluation Ready approach highlights data planning as an integral part of evaluation planning (DIIS 2019a). The evaluation planning process involves developing a program logic, an evaluation strategy and a data matrix (box 9.8).

Box 9.8 A data matrix to help with evaluation planning

A data matrix is a tool for organising evaluation questions and sub-questions, and identifying the information needed to address them. A data matrix should inform ongoing data collection requirements, highlight existing data sources that could be used during evaluation, and guide any additional data collection during evaluation. It should document the evaluation questions, indicators (what is measured), metrics (how it is measured), and any contextual/comparison data, as well as assign responsibility for data collection.

The Department of Industry, Innovation and Science's data planning requirements cover:

- developing evaluation questions from the program logic
- identifying performance indicators and data sources
- articulating data collection responsibilities and timeframes
- identifying the required data, including the limitations of particular sources.

The following is a high-level example of a data matrix for the Job Commitment Bonus (JCB) for Young Australians evaluation by the former Department of Jobs and Small Business. The JCB is a payment to young people to encourage them to find work.

		Evaluation questions			
		Did the JCB influence take up and retention of employment among participants?	Did employment outcomes increase following the introduction of the JCB?	What other types of financial incentive or support, if any, were received in conjunction with the JCB?	Was the JCB cost effective?
Data Sources	Administrative data (DJSB)	✓	✓	✓	✓
	Administrative data (DHS)	✓			✓
	Employer survey	✓			
	Employment service provider survey	✓			
	Focus group	✓			
	Post program monitoring survey		✓		
	Modelled estimates				✓

Sources: DIIS (2019a); DJSB (2019).

Data planning should involve policy or program managers, evaluators, and those that are affected by the policy or program. And when considering what data are needed for evaluation, it is important to think about the full range of information that is needed to answer the evaluation questions, including:

- *baseline or control group data* — descriptions of the nature of the policy or program and how it is to be implemented (and, later, how it was implemented)
- *participant characteristics* — the characteristics of those affected by the policy or program, especially characteristics that the program logic suggests will be relevant
- *factors that affect the policy or program* — other interventions, events or circumstances that could affect the policy or program
- *outcomes and impacts (short, medium and long term)* — outcomes that are expected from the theory of change, and other data that could explain why they do or do not occur
- *unintended/unexpected effects* — outcomes that are not anticipated and data that may explain why they occur
- *appropriateness and relevance of the policy or program* — whether the policy or program addresses participant and community needs
- *other considerations* — these may come from the theory of change, arise later through initial evaluation activities (for example, after a first round of engagement) or seek to address concerns from other interested parties (adapted from Perrin 2012).

Data planning should also consider how data are best collected, including who should collect them, when they should be collected, and whether data collection methods are suitable for participants (box 9.9). Consideration should also be given to any potential issues that may affect the evaluation, such as ensuring representative samples, data burdens on participants, and access or approvals to use existing data (discussed further below).

Box 9.9 An example of adapting data collection processes for local contexts

Families and Schools Together (FAST) is an early intervention program designed to strengthen family functioning and build protective factors for children. It is an international program that was adapted to the Northern Territory context.

The program had already established evaluation processes, but when preparing to evaluate the adapted program it became clear that the evaluation tools would not work in remote communities, as the language and concepts that the tools used were not always meaningful to program participants.

The new evaluation processes needed to be culturally relevant and meaningful to participants, while still being rigorous and consistent with the international evaluation tools. A psychometric survey tool was not producing reliable results and was replaced with a narrative inquiry tool that used pictures and symbols that were meaningful to local participants, but also could be translated into definitive outcomes.

The process took time and effort and required several trials of the evaluation tools to ensure that they worked.

Source: Muir and Dean (2017).

Weighing up the costs and benefits of collecting data

As discussed in section 9.2, collecting data is costly for government, evaluators and participants. Data planning should carefully consider what data are necessary for an evaluation, taking into account the costs of data collection relative to its value.

For example, when planning to collect data on outcomes for an evaluation, the following questions should be considered.

- Is the outcome or measure important to those affected by the policy or program?
- Does the outcome align with the intended objectives of the program? Is this a core objective of the program?
- If the outcome is not a key goal of a program, is it a significant unintended outcome? Evaluations should consider unintended outcomes where they may be influential in deciding the benefits and impacts of a program.
- Is the outcome within the program's sphere of influence? Is it reasonable to expect the program would have influenced this outcome?
- Will the program be at the right stage of delivery to produce the particular outcome? Outcomes should be achievable within the timelines of the evaluation. For example, a long-term outcome should not be assessed immediately after the end of the program.
- Can the outcome be measured? Evaluations should consider whether the indicators for a particular outcome can be reliably measured. Where outcomes are deemed important but difficult to measure reliably, limitations should be acknowledged.

-
- Will measuring the outcome provide useful information about whether the program is effective or not? Evaluation findings should help to make decisions about the program. If measuring an outcome gives interesting, but not useful, information, it is probably not a priority (adapted from AIFS 2017).

Quantitative or qualitative data?

Another consideration during the planning phase is the extent to which quantitative or qualitative data (or some combination of the two) should be used in an evaluation. Quantitative and qualitative data both have strengths and weaknesses (table 9.1). Most evaluations use both types of data, taking advantage of the strengths of each. As discussed in chapter 4, a mixed approach can overcome some of the limitations of only using one type of data. Some examples of how quantitative and qualitative data can complement each other include:

- using qualitative data to identify issues or obtain information about variables that cannot be obtained through quantitative approaches, or vice versa
- generating hypotheses from qualitative data to be tested using quantitative data (for example, by using qualitative data to identify subgroups that may experience different impacts from a policy or program, then exploring the nature of the impacts for these subgroups using quantitative data)
- using qualitative data to understand unanticipated results identified using quantitative data, or vice versa
- using qualitative data to verify or triangulate results from quantitative data, or vice versa (adapted from Greene, Caracelli and Graham 1989).

Table 9.1 Strengths and weaknesses of quantitative and qualitative data

	<i>Strengths</i>	<i>Weaknesses</i>
<i>Quantitative</i>	• Study findings are more generalisable • Samples can be selected to be representative • Estimates can speak to the magnitude and distribution of impacts, as well as costs and benefits • Estimates can be more easily compared and replicated elsewhere	• Much information cannot be collected through quantitative survey instruments • Self-reported information may be inaccurate or incomplete • There is a lack of contextual or other information to explain variation • Quantitative data collection can be expensive • Research methods may be more rigid and not allow for much flexibility • Correlations may not be causal • Incorrect hypotheses may lead to incorrect conclusions • Errors in sampling may produce inaccurate conclusions
<i>Qualitative</i>	• More flexibility for the evaluation design to change • Data collection can be more easily targeted and adapted for certain groups • Greater scope to focus on the 'whole picture' • Can provide a more nuanced understanding of complex situations or behaviour • More accessible to non-technical audiences	• Not always generalisable • Different perspectives may confuse the 'takeaway' message or conclusion • May be more subjective • Harder to assess validity without careful documentation

Sources: Adapted from Bamberger (2012a, 2012b).

Addressing sampling, non-response and attrition

Data planning should take into account sampling needs for the evaluation. As discussed above, this is important for mainstream policies and programs, as Aboriginal and Torres Strait Islander people are typically a small percentage of participants in mainstream policies and programs. A standard sampling technique may not result in enough respondents to produce meaningful results. Attrition and non-response can also significantly reduce sample sizes, with consequences for the methodological approaches that can be undertaken and/or the robustness of results. Sampling strategies are discussed in chapter 5.

Attrition and non-response can be particularly problematic where they do not occur evenly across the sample and can lead to certain groups being underrepresented. When this happens, findings may not reflect the true impact of a policy or program, or may limit how reliably any results can be generalised to the whole population. As discussed above, various factors — including language, Internet access, literacy and high geographic mobility — can be

barriers to participating in data collection that may lead to attrition or non-response (PC 2020, p. 212).

When planning for data collection, agencies should consider how they choose their sample to ensure that samples:

- are as representative of the wider population as possible
- can be disaggregated along relevant dimensions such as age, gender and remoteness
- are free from bias to the greatest extent possible.

Engaging with Aboriginal and Torres Strait Islander people on data planning and management

There was a widespread view among participants that embedding Aboriginal and Torres Strait Islander leadership in data governance was important for the Indigenous Evaluation Strategy (the Strategy). For example, the National Aboriginal Community Controlled Health Organisation said:

Great care needs to be taken in developing data governance arrangements in the new [Indigenous Evaluation Strategy]. It is essential for data governance principles and structures to recognise and accommodate the need for Aboriginal and Torres Strait Islander peoples' self-determination and community control in all steps of collecting and using data. In practice, this means ensuring the collection and reporting of data pertaining to Aboriginal and Torres Strait Islander people is culturally appropriate, safe, competent and in the best interests of Aboriginal and Torres Strait Islander peoples. (sub. 95, p. 6)

AbSec – NSW Child, Family and Community Peak Aboriginal Corporation said that greater Aboriginal and Torres Strait Islander engagement in data collection, use and governance would improve the relevance of data and empower Aboriginal and Torres Strait Islander decision making:

By empowering Aboriginal communities with relevant data, and the opportunity to define, collect and analyse such data in accordance with our own perspective, priorities and aspirations, Aboriginal communities will be better placed to make informed decisions about policies and programs, and to determine and administer systems for our social, cultural and economic development. (sub. 9, p. 7)

Many also noted that greater Indigenous data governance could help to improve the quality of evaluations. Sophia Couzos, for example, said that Indigenous data governance could help establish trust:

Importantly, data ownership (governance) issues are critical to establish trust in evaluation methodologies. The Indigenous Evaluation Strategy should identify Aboriginal data governance protocols as a priority, but these are currently lacking. (sub. 92, p. 3)

Social Ventures Australia argued that greater Aboriginal and Torres Strait Islander engagement on data could inform culturally competent data collection methods and ultimately improve data quality:

The data collection methods used — how people are consulted, how much time is spent visiting communities and when, the languages and methods used and who is consulted — are all factors that influence the quality of data, and the potential for evaluation to support community capacity and self-determination. (sub. 83, p. 11)

How can Indigenous data governance be applied in practice?

There is no one approach to Indigenous data governance. Approaches may differ across mainstream and Indigenous-specific datasets, and will depend on the context, use and content of the data. A range of different approaches have been applied in practice in Australia and around the world (box 9.10).

The Organisation of Economic Cooperation and Development recently recommended greater Indigenous data governance to improve regional development through:

- including Indigenous representatives in the governance of national statistical agencies to provide advice on strategic and operational issues impacting on Indigenous peoples (for example by providing advice on statistical definitions, the design of wellbeing indicators and data collection methods)
- implementing protocols and agreements to enable the pooling of data between different agencies to increase sample sizes and the availability of data
- adapting data collection methods to the needs of Indigenous peoples through interview-administered surveys in Indigenous languages that include communities in the data collection process
- providing tools and capabilities for Indigenous organisations to collect their own data on issues that are important to their communities, and support more informed decision making about development (OECD 2019, p. 53).

Box 9.10 **Approaches to Indigenous data governance**

Longitudinal Study of Indigenous Children (LSIC) (Australia)

LSIC is a longitudinal study of Aboriginal and Torres Strait Islander children with 10 waves. It is funded by the Department of Social Services and overseen by a Steering Committee with significant Aboriginal and Torres Strait Islander representation. To access data, researchers need to comply with three principles:

- acknowledge ‘standpoint’ when applying for and reporting LSIC data — researchers need to acknowledge their institutional context as well as personal context, such as cultural background, work background and depth of experience working with Aboriginal and Torres Strait Islander people
- cultural competency — researchers need to respect, understand and acknowledge the benefits, values and realities of Aboriginal and Torres Strait Islander people and communities
- contributing knowledge to build better outcomes — sharing findings with communities to be accountable to them and ensuring better outcomes for them.

First Nations Regional Health Survey (Canada)

The First Nations Regional Health Survey was first developed in the 1990s. It was designed, implemented and disseminated by participating First Nations people. Today, it is run by the First Nations Information Governance Centre, which was originally a committee of the Assembly of First Nations — the First Nations peak body — but was incorporated in 2010. Ten regional partners provide oversight and the survey cannot be undertaken without a resolution from community leaders consenting to the survey. The survey is funded by Health Canada, who access data at a national level via a licence-to-use agreement, but ultimate ownership lies with the First Nations.

First Nations Health Authority Tripartite Data Quality and Sharing Agreement: British Columbia (Canada)

The First Nations Health Authority, along with the government of British Columbia and the Canadian federal government, signed the Tripartite Data Quality and Sharing Agreement in 2010. It sought to improve the quality, sharing and use of First Nations health data. The agreement is managed by a committee with three voting members, each representing a party to the agreement. The committee works mainly by consensus.

Albuquerque Area Southwest Tribal Epidemiology Center (AASTECC) (United States)

AASTECC, established in 2006, is one of 12 Tribal Epidemiology Centers in the United States (see appendix D for more information on the role of Tribal Epidemiology Centers). The Center provides a range of epidemiology services to 27 tribes, which includes program evaluation. It works with tribes to assess and ascertain their data needs, collect data and analyse data. The Center also acts as an intermediary to share data and work with government entities to improve data quality.

Whānau Ora (Aotearoa/New Zealand)

Whānau Ora is a devolved commissioning approach to service delivery. The Whānau Ora partnership group provides strategic direction and consists of representatives from the New Zealand Government and the Iwi Chairs Forum. A key part of its role is to establish an outcomes framework and to monitor progress against it.

Sources: Bruhn (2014); DSS (2019c, 2019b); NCLD (2018); TPK (2015); Walter and Suina (2019).

When data are collected by or for government, implementing Indigenous data governance requires flexibility, trust, commitment and technical capability from all parties. Bruhn (2014) examined several examples of Indigenous data governance from Australia, Canada and the United States and found some common themes and challenges.

- In negotiations of data sharing arrangements between governments and Indigenous organisations, trust-building and frequent contact between the parties are important in addressing the legal, policy, and capacity challenges that arise.
- At the operational level, both commitment and technical capacity are required to implement data linkages; such capacity is often in short supply on both government and Indigenous sides.
- Some Indigenous data governance initiatives start small and build up slowly, seeking support from senior levels as they progressed; others are initiated through high-level political accords that are then progressively implemented.
- Indigenous data sharing arrangements vary in their interpretation and embodiment of Indigenous data governance principles. Indigenous partners support approaches that allow for compromise, as long as they are assured that the spirit and intent of the principles are preserved.
- Some examples feature local, regional, and even national-level stewardship of Indigenous data. Other initiatives are tightly focused on the community level. That said, staff at all organisations interviewed noted that their priority was to ensure that communities gain access to, and have a voice in, the governance of the data concerning them.

Indigenous data governance in the Indigenous Evaluation Strategy

Indigenous data governance should be embedded in the Strategy as a means of:

- ensuring that collected data reflect the priorities of Aboriginal and Torres Strait Islander people and communities
- improving trust in data collection and governance arrangements
- ensuring that data are being collected, managed, used and stored in an ethical and culturally appropriate way
- building on existing Aboriginal and Torres Strait Islander data capability.

As with Indigenous data governance more broadly there are many ways evaluations can build in Indigenous data governance (box 9.11). These include:

- Aboriginal and Torres Strait Islander people being engaged early in an evaluation, including being involved in defining evaluation questions and identifying the data required to answer the questions
- Aboriginal and Torres Strait Islander people developing and/or agreeing to data collection procedures

-
- the use and further availability of data (including sharing with third parties) being negotiated with Aboriginal and Torres Strait Islander people
 - Aboriginal and Torres Strait Islander people working with the evaluation team to interpret and validate data, and endorse the release of information.

Box 9.11 Indigenous data governance in an evaluation

Stronger Communities for Children Program

Evaluation methods were developed with the community by a local research team. Local researchers helped recruit participants, conduct interviews and analyse data. Communities were invited to give feedback on the program logic and findings. A short report was provided to each community involved in the case study. (Winangali Ipsos Consortium and IPSOS 2017)

Maranguka Justice Reinvestment Project

This program is a community-led, data-driven approach to justice reinvestment. Local people are trained in data collection methods. Local governance groups regularly receive data and participate in annual reflection workshops to make sense of the data and reflect on progress. The evaluation used data collected and collated during program design and delivery to assess impact. (Just Reinvest NSW, sub. 80; KPMG 2018)

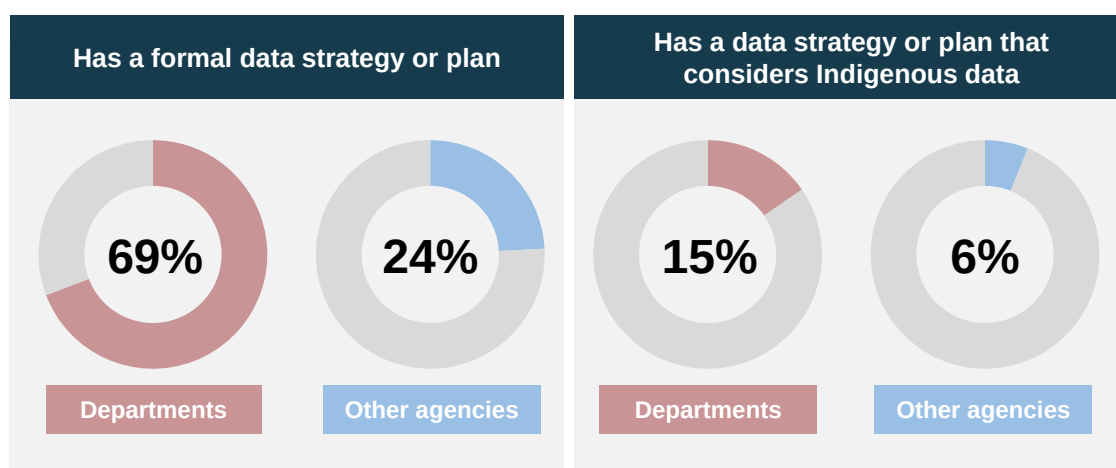
Indigenous Australians Health Program

This evaluation was designed with a steering group of which more than half of members were Aboriginal and Torres Strait Islander people. In the evaluation itself, which is currently underway, 20 sites will be examined, with each site able to articulate their preferred level of data collection. The first year of the evaluation will produce a data feasibility study, where evaluators will work with participants to identify potential data sources and assess their suitability for the evaluation. Throughout the evaluation, each site will be provided with regular reports to support planning and continuous improvement. There will be ongoing sense-making workshops with evaluation participants. (Department of Health, sub. 35, attachment A; Baily et al. 2018)

Indigenous data governance is an important mechanism to ensure that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred. However, in the context of mainstream policies and programs, the extent to which these are built into evaluation governance arrangements would need to be balanced with the scope and size of the evaluation, its impact on Aboriginal and Torres Strait Islander people, and privacy and confidentiality concerns.

Agencies should also consider how best to engage with Aboriginal and Torres Strait Islander people in their general data planning activities. Many agencies, particularly departments, have started to strengthen their own data capabilities by developing data registries, building data skills, and creating data strategies and plans (figure 9.3). These efforts should support better data for evaluation by improving consistency, interoperability, and discoverability of data, as well as building agency culture and capability to collect and use data.

Figure 9.3 Data planning is relatively widespread in departments, but few agencies consider Indigenous data in their data plans
Percentage of Australian Government agencies^{a,b,c}



^a Based on a sample of 46 Australian Government agencies. ^b Total for agencies that report having a formal data strategy or plan excludes agencies that reported having a privacy policy but no other data strategies or plans. ^c Total for agencies with a data strategy or plan that considers Indigenous data includes agencies with a general data strategy or plan that mentions Aboriginal and Torres Strait Islander data issues, and those with a separate Indigenous data strategy or plan.

Source: Productivity Commission Indigenous Evaluation Strategy information request, appendix B.

However, few agencies have any formal policy, strategy or guidelines for collecting and managing information or data about Aboriginal and Torres Strait Islander people. Only 15 per cent of departments and very few other agencies consider Indigenous data as part of their general data planning activities or have specific Indigenous data strategies or plans (figure 9.3). Some examples of agencies that do consider Indigenous data in their data planning are described in box 9.12. In addition, several agencies told the Commission that they use the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) Guidelines for Ethical Research in Australian Indigenous Studies (2012) to guide them when collecting, using and managing data about Aboriginal and Torres Strait Islander people.

Box 9.12 **Planning for Indigenous data in Australian Government agencies**

Former Department of the Environment and Energy

The Department of the Environment and Energy's information strategy outlines key actions to improve the Department's approach to Aboriginal and Torres Strait Islander data. It commits the agency to:

- developing a departmental Indigenous data standard
- working with the Department's Reconciliation Action Committee to further understand Indigenous data throughout the Department
- including appropriate markers to identify Indigenous data in the Department's dataset register
- explicitly considering Indigenous data for all improvements to data related policies, procedures and technology (sub. 101, p. 5).

Department of Health — Indigenous Health Division

The Indigenous Health Division's data strategy guides the use, improvement and development of data. The Strategy is informed by the widely accepted Australian Health Ministers' Advisory Council (2006) National Aboriginal and Torres Strait Islander Health Data Principles, and includes:

- ensuring respectful engagement and partnership with Aboriginal and Torres Strait Islander people — including working through relevant health committees and recognising the role of culture in the maintenance of health
- ensuring the cultural safety of data collection and use — including that data management protocols are followed and ethics clearance obtained where necessary
- adopting a strengths-based approach in data development where appropriate — including celebrating health gains and exploring protective health factors
- maximising the value of data collection — including improving the use and quality of existing data, ensuring that benefits of data outweigh collection burdens, and working effectively with data limitations (for example using data linkage to overcome them)
- ensuring access to information and data — making effective use of internal data infrastructure such as the enterprise data warehouse, and external expertise from the Australian Bureau of Statistics and Australian Institute of Health and Welfare
- enhancing data analytic capability and capacity — developing capability within the division and leveraging expertise in other divisions
- strategising data investment — assessing current data availability and identifying data development priorities (pers. comm., 25 May 2020).

To improve Indigenous data governance, agencies should engage with Aboriginal and Torres Strait Islander people as part of their general data planning activities about how data might be best collected and used, including to effectively undertake evaluations. Agencies should address:

- how Indigenous data governance can be strengthened
- the use of appropriate standards, including the Indigenous status standard
- whether there are common outcomes across the agency that might be measured

-
- appropriate data sharing and release protocols
 - ethical and culturally appropriate data collection methods.

By engaging with Aboriginal and Torres Strait Islander people during data planning, agencies will be able to develop and identify solutions earlier on and across the agency, rather than having to address similar concerns in each evaluation.

Developing a data dictionary and guidance

One of the challenges with collecting data that are more relevant for Aboriginal and Torres Strait Islander people is developing indicators that adequately capture what is important for their lives (Cairney et al. 2017). Kelaher et al. commenting on the challenges of evaluating outcomes that reflect a holistic definition of health, said:

The discrepancy between program and evaluation outcome measures reflects the lack of widely accepted and validated measures for indicators of holistic health, and [the comparatively greater] ease and precision in measuring biomedical states of mortality and morbidity. This reflects a real need for measures that accurately capture the holistic concept of health. (2018b, p. 15)

Indicators can be costly and time-consuming to develop, particularly where they need to apply across a range of contexts and cultures. For example, the Mayi Kuwayu survey (box 9.4) was developed over three years with 24 focus groups and two pilot ‘proof-of-concept’ studies (Jones et al. 2018).

While some of the indicators that are relevant for answering evaluation questions are likely to differ across evaluations, locations, programs and contexts, there are potentially some core indicators of Aboriginal and Torres Strait Islander people’s wellbeing that could be used for multiple evaluations.

Given the costs and complexities of developing new indicators, there is merit in establishing a standardised, whole-of-government data dictionary that agencies can draw on when developing indicators or deciding on the types of indicators that they will collect for monitoring or evaluation purposes. Such a dictionary would not replace the need to engage with Aboriginal and Torres Strait Islander people on the right questions to ask, and the right data to collect, when planning each individual evaluation. However, a data dictionary would foster the better development and use of indicators by:

- ensuring consistency and comparability of data content and definitions across government for some core indicators of Aboriginal and Torres Strait Islander people’s wellbeing
- avoiding data duplication and differing definitions for indicators of similar outcomes
- reducing costs and burdens of developing data collection tools, both on agencies and Aboriginal and Torres Strait Islander people.

The AIHW is well placed to undertake this task, having prior experience in indicator development and Indigenous data. The AIHW should not collect data for evaluations —

rather it should develop the metadata and procedures that agencies can use to monitor outcomes of significance to Aboriginal and Torres Strait Islander people in a consistent way.

In developing the data dictionary, the AIHW should partner with Aboriginal and Torres Strait Islander people, with consideration given to:

- how to best use existing Aboriginal and Torres Strait Islander data expertise, including the Indigenous Data Network, Maïam nayri Wingara, Aboriginal and Torres Strait Islander academics, and Aboriginal and Torres Strait Islander organisations
- what outcomes and indicators will be informative to government as well as Aboriginal and Torres Strait Islander people
- standards that already exist (such as the Indigenous status standard)
- the evaluation priorities under the Strategy (chapter 6).

The data dictionary should also include guidance on how to collect and use data, much like the best practice guidelines that complement the Indigenous status standard. This guidance may include how to ensure appropriate consent is granted, how to collect data in a culturally safe manner, and advice on when data sharing and release may lead to harm to the community.

INFORMATION REQUEST 9.1

The Commission is seeking feedback on the indicators that are highest priority for inclusion in the data dictionary. What outcomes and indicators are most informative for government and Aboriginal and Torres Strait Islander people?

Making better use of existing data

Benefits of using existing data in evaluation

There is scope for Australian Government agencies to make greater use of existing data sources in evaluation. There will often be a need for data collection during evaluation, as not all evaluation questions can be answered with existing data. However, there are potentially many benefits of using existing data sources for evaluation — to the greatest extent possible — rather than collecting data as part of the evaluation process. These benefits range from reducing the costs of evaluation to enhancing rigour, and include:

- *reduced participant burden* — participants are not asked to provide information multiple times on the same topics
- *increasing the value of existing data* — making greater use in evaluation of datasets that have been collected for other purposes increases their value
- *lower cost and time burden for evaluators* — while there may be some cost and delay to accessing existing datasets, it is likely to be smaller than undertaking primary data

collection for an evaluation. Using existing data also reduces the need to have evaluators with expertise in survey or interview techniques

- *large sample sizes/wide coverage* — many existing data sources (for example, the census and departmental administrative data) have close to universal coverage of populations of interest. This is particularly valuable for evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, who are a small and geographically-dispersed population, and for whom generating a sufficient sample size through a survey may be costly. Wide coverage also allows for the identification of reliable comparison groups who may be hard to contact or target in surveys
- *longitudinal nature of data* — existing longitudinal data allow for comparison of outcomes before and after program participation, and for longer-term outcomes to be assessed
- *increasing accuracy* — data that are collected passively or for other purposes may be less prone to enumerator or social desirability bias than those collected as part of an evaluation or during program participation. Data collected contemporaneously over a long time period may also be less prone to measurement error due to recall bias than one-off surveys that ask respondents about their situation at several points in time (Feeney et al. 2015).

One way to increase the use of existing data in evaluations is by improving the quality of data collected for monitoring (including through early planning and engaging with Aboriginal and Torres Strait Islander people, as discussed above). Another is to improve data sharing and data linkage arrangements.

Data sharing

Data sharing is the process of making data available to third parties. Data sharing generally involves constraints on the use of the data, the timeframe for use of the data, and/or the technology used to analyse it. The challenges of data sharing are discussed in detail in the Commission's *Data Availability and Use* inquiry report. They include:

- *fragile community understanding and trust* — including community concern over privacy breaches and re-identification of data
- *complex legislative arrangements* — including inconsistent and unclear applications of privacy laws and/or secrecy provisions
- *risk aversion* — including that data sharing is often restricted, or approval processes are long and complex due to concern about the potential downsides of sharing data
- *lack of leadership* — including a lack of empowerment and consistency in sharing data (PC 2017a).

In response to the recommendations of the *Data Availability and Use* inquiry, the Australian Government established the Office of the National Data Commissioner (ONDC) to:

- promote greater use of public sector data

-
- drive innovation and economic benefits from greater use of public sector data
 - build trust within the Australian community around the Australian Government's use of data (ONDC 2019).

The ONDC is currently developing a data sharing framework that will facilitate the sharing of public sector data (box 9.13). The framework will be underpinned by new data sharing legislation. Under the new legislation, data may be shared for evaluations of government policies and programs, both between agencies and with external researchers.

Box 9.13 Main features of the proposed data sharing and release regulatory framework

In response to the Commission's report into *Data Availability and Use* (PC 2017a), the Australian Government agreed to develop legislation to enable the greater sharing of public sector data. The draft legislation was set to be released for consultation in mid-2020. The legislation will not seek to compel data custodians to release data. Rather, it will provide data custodians with a pathway to share data that they may not have had otherwise.

The new legislation will:

- provide Australian Government agencies or other data custodians with an alternative pathway to share public sector data to accredited entities such as government agencies, state and territory authorities, and non-government entities such as universities
- take a principles-based approach to safeguarding the sharing of public sector data, notably using the 'Five Safes Framework', which takes a multidimensional approach to managing disclosure risk
- empower the National Data Commissioner to develop requirements and guidance to support government agencies and data custodians
- establish clear governance arrangements, including enforcement and accountability mechanisms.

Source: DPMC (2019c).

These reforms are a good first step in improving data sharing and increasing data use. They will, however, need to be complemented by efforts by the ONDC to promote data sharing and guide agencies to better share their data. Further investments in sharing infrastructure and other data improvements are also needed. For example, the Independent Review of the Australian Public Service (APS) commented:

The Government must build on this momentum by further strengthening data and analytics capabilities, embedding the use of data in ways of working and establishing technical governance and whole-of-government data infrastructure to support seamless data-sharing between agencies, trusted partners and Australians. (DPMC 2019e, p. 178)

Data linkage

Data linkage can be a powerful tool for use in evaluations. Data linkage involves linking multiple datasets together by using features common to these datasets (box 9.14). Data linkage results in richer datasets that can be used to gain insights that are not possible with a single dataset or survey.

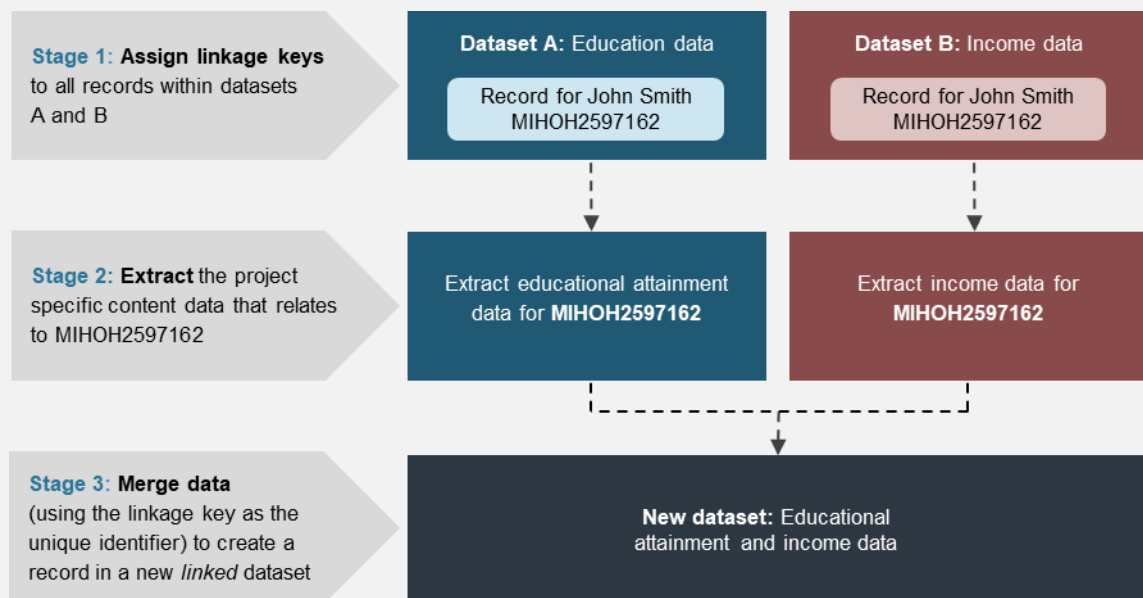
Box 9.14 The process of data linkage

Data linkage creates new datasets by combining existing data. This can have substantial advantages, both in terms of making better use of existing data and discovering new patterns in the data.

There are two main techniques for linking datasets: deterministic and probabilistic. In deterministic linkage, the datasets to be linked contain a unique identifier for each entity (for example, a tax file number). The linkage matches the unique identifier across the datasets.

Probabilistic linkage is based on calculating the likelihood that two records match when there is no unique identifier. Probabilistic linking is a slower process but allows data to be matched from a wider array of sources.

Where there are no unique identifiers, a linkage key can be created. The linkage key is a code created from identifiable information, such as names or addresses, that are included in both datasets to be linked. These linkage keys can be saved and reused across projects. Using linkage keys prevents the need to share identifiable information with researchers, and therefore contributes to minimising the risk of privacy breaches. An example is shown below.



Source: PC (2017a, p. 427).

Many participants noted the value of linked datasets for evaluation. For example, Empowered Communities commented on their value in obtaining baseline datasets:

These are rocks that have been turned over many times but we are not going to be able to build effective baseline data sets to underpin evaluation processes without addressing some of the data gap and quality issues, or without safely exploring alternative means of building pictures of this population level data and this may include considering data linkage and data integration processes. (sub. 41, p. 8)

Data linkage can also be used to track longer-term outcomes. The AIHW commented:

Some outcomes cannot be observed while a program is operating as they require generational change. ... It may be possible to observe the subsequent impact of programs and policies through data linkage and it could be possible to assess whether particular types of assistance have a positive impact on key outcomes compared with a control group. (sub. 99, p. 5)

Data linkage could also be useful where:

- mainstream datasets do not collect data on outcomes that are of particular relevance to Aboriginal and Torres Strait Islander people. Linking data on Indigenous-specific outcomes to a mainstream dataset allows these outcomes to be included in an evaluation of a mainstream program
- sampling strategies are unlikely to capture a sufficiently large sample of Aboriginal and Torres Strait Islander people. Linking administrative data means a much larger sample — and potentially the entire population — can be assembled
- data have not been collected for a control group, or are not available once participants leave a program. Linking allows for outcomes to be measured for them as well as program participants.

Two important risks need to be considered when undertaking data linkage: ensuring adequate consent has been obtained for linkage, and the potential for re-identification.

Data may be collected for a single purpose without participants explicitly consenting to share them more widely or for them to be used in research and evaluation (PC 2016b, p. 157). In such circumstances, evaluators and agencies need to carefully consider what consent may be needed to ethically use linked data, and should seek guidance from a Human Research Ethics Committee where appropriate.

While data linkage usually uses de-identified data, there may be concerns that data linkage could enable the re-identification of de-identified data. As more datasets become available, re-identification is potentially more likely if data are not carefully managed (PC 2017a, pp. 156–157). These risks, however, can be managed so that data sharing and data linkage are still feasible. The most common approach to managing these risks is through the Five Safes Framework (box 9.15) that will be adopted under the data sharing legislation and is also used by the ABS and AIHW for their data sharing arrangements.

Box 9.15 **The Five Safes Framework**

The Five Safes Framework is an approach to managing risks of re-identification and breaches of confidentiality. It acknowledges that there will always be some risk, but with appropriate safeguards, this risk can be managed while maximising the benefits that can be obtained from the data.

The Five Safes Framework focuses on five orthogonal risk axes.

- Safe people: Can the researchers be trusted?
- Safe projects: Is the purpose of use appropriate? What analysis is being done?
- Safe settings: Does the access environment prevent unauthorised use?
- Safe data: Can the data disclose identity?
- Safe outputs: Are the statistical results non-disclosive?

More recent approaches using this framework have been based on the realisation that, if one or two of the axes introduce higher risk, the overall risk of disclosure may still be low as there are multiple ways in which risk can be managed. In practice, this has resulted in virtual laboratories or trusted access models that provide riskier data in a safer environment. For example, to access some de-identified confidential data at the Australian Bureau of Statistics, users must log on to DataLab, a remote access environment, where data cannot be digitally exported and any outputs require manual inspection for risk of statistical disclosure.

Source: PC (2017a, p. 185).

The Australian, state and territory governments have invested significantly in enduring integrated datasets — one example is the Data Integration Partnership for Australia (DIPA) (box 9.16). These assets can be useful for conducting evaluations and some projects are already using them. For example:

- the former Department of the Environment and Energy used the Multi-Agency Data Integration Partnership (MADIP) — an integrated dataset linking together census, tax, and other individual data — to measure the impact of Aboriginal and Torres Strait Islander environmental programs (ABS 2019)
- the Department of Industry, Science, Energy and Resources is using the Business Longitudinal Analysis Data Environment (BLADE) — a data linkage environment that links together business information by Australian Business Number — to evaluate its programs. The Department works with the Australian Bureau of Statistics to integrate its program data into BLADE. The use of BLADE allows for a longer-term assessment of outcomes, as well as the identification of appropriate counterfactuals (DISER 2017).

Box 9.16 **Data Integration Partnership for Australia**

The Data Integration Partnership for Australia (DIPA), started in July 2017, is a three-year \$130.8 million investment aimed at improving the use and quality of the Australian Government's data assets. There are over 20 Commonwealth agencies involved in DIPA.

DIPA has several components. These include:

- *Data integration* delivered by Australian Bureau of Statistics and Australian Institute of Health and Welfare. These two bodies deliver the core technical infrastructure to support DIPA by providing tools to integrate and link data, including the production of enduring longitudinal and integrated data assets. These include the Multi-Agency Data Integration Partnership, Business Longitudinal Analysis Data Environment and National Integrated Health Services Information Analysis Asset.
- *Data assets* including:
 - *Data Exchange*, delivered by the Department of Social Services, which seeks to streamline data collection for client-facing government funded programs
 - *National Education Evidence Base*, delivered by the Department of Education, which brings together multiple sources of education data including tertiary and early childhood data.
- *Data analytical units* use integrated data to solve complex policy issues that cross over multiple portfolios and feed this analysis into policy delivery processes and improved service delivery. There are five units: the Social Health and Welfare Analytical Unit; the Economic Data and Analysis Network; the Physical Environment Analysis Network; the Government Business Analytical Unit; and the Central Analytics Hub.
- *Better communication and engagement* about data initiatives by the Department of the Prime Minister and Cabinet.
- *Technical review and advice* provided by Data61 in the Commonwealth Scientific and Industrial Research Organisation.

While DIPA does not have an explicit focus on evaluation, many of these data improvement activities have the potential to be useful for evaluation.

Source: DPMC (2019b).

Further improvements to DIPA datasets could improve their relevance for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. For example, Indigenous-specific survey data, such as the National Aboriginal and Torres Strait Islander Social and Health Surveys, could be integrated into MADIP.

Another potentially useful project for evaluation is the Department of Social Services' Data Exchange (DEX). DEX is a performance reporting tool for funding agencies and service providers. It aims to streamline data collection by standardising data collections across multiple agencies, including state and territory government agencies. Service providers are able to receive reports about their performance (DSS 2019a). It also has the potential to enable data linkage across multiple programs. However, there are some limitations of DEX, which include:

- *data collection on certain outcomes is not mandatory* — evaluations that have tried to use data from DEX have found many missing observations on key variables, so that there

is insufficient coverage for the data to be useful. The Commission understands that there has been a recent change to mandatory collection and now a wider range of variables are collected mandatorily

- *qualitative data cannot be collected via DEX* — qualitative data are often used in evaluation. Allowing service providers to collect qualitative data would increase DEX's usefulness for evaluation
- *there are few opportunities for input into what is collected* — DEX variables are determined by the Department of Social Services, with limited opportunities for engagement with Aboriginal and Torres Strait Islander people or other potential data users on what data can and should be collected.

DIPA is currently funded until 2020. The Australian Government's response to the Independent Review of the APS noted the role of DIPA in improving APS capability to interrogate and use insights from data, but the Government has not yet committed any additional funding towards it. It is not clear to what extent DIPA will continue beyond 2020 if it is not further funded. Continued funding for data linkage projects will be important for improving linked data for evaluation under the Strategy.

INFORMATION REQUEST 9.2

Which data linkages would best support the Indigenous Evaluation Strategy? Which data linkages are important to Aboriginal and Torres Strait Islander people?

10 Monitoring and reviewing the Strategy

Key points

- Monitoring arrangements for the Indigenous Evaluation Strategy (the Strategy) have two important functions. These are to monitor:
 - the performance of Australian Government agencies against the Strategy
 - how effective the Strategy has been in encouraging high-quality and useful evaluations.
- Responsibility for these monitoring functions should be housed with the Office of Indigenous Policy Evaluation (chapter 7). Regular, public reports should be produced that outline findings and recommendations based on the monitoring of the Strategy's implementation. The first report should be released two years after the Strategy is endorsed, then biennially thereafter. These reports should:
 - report on the extent that Australian Government agencies have complied with the Strategy
 - identify good and/or innovative practice in the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people
 - provide an assessment of the extent to which the Strategy has been effective in encouraging high-quality and useful evaluations
 - formalise evaluation priorities, including identifying areas for cross-agency evaluations
 - provide recommendations on how the implementation of the Strategy could be improved.
- The Australian Government should commission an independent review of the Strategy five years after it is first endorsed. Reviewing the Strategy is an opportunity to assess:
 - whether the Strategy's principles remain fit-for-purpose, and if not, in what ways they should be updated
 - the extent to which the Strategy has been effective in encouraging higher quality and more useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, including the extent that the key actions of the Strategy have been effective
 - the performance of those tasked with overseeing the Strategy and whether changes to the policy environment require changes to oversight arrangements.

This chapter examines monitoring (section 10.1) and review (section 10.2) arrangements for the Indigenous Evaluation Strategy (the Strategy). The chapter looks at why these functions are important and what they need to achieve, and outlines a proposed approach for monitoring and reviewing the Strategy.

10.1 Monitoring arrangements for the Strategy

What do monitoring arrangements need to achieve?

The monitoring arrangements for the Strategy need to:

- monitor the performance of Australian Government agencies against the Strategy
- monitor how effective the Strategy is at encouraging high-quality and useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

Monitoring the performance of agencies against the Strategy

The Strategy provides principles-based guidance for Australian Government agencies on what good quality looks like for all stages of evaluation — planning, conduct, reporting and use. If the Strategy is to raise the bar on the quality and usefulness of evaluations, agencies will need to follow this guidance.

While agencies have some incentive to evaluate their policies and programs (not least because evaluation can identify ways to make these more effective and efficient), evaluation does not always happen in practice, and when it does, it can be done more for compliance reasons rather than finding out whether or not policies or programs are working or how they can be improved. Without external oversight, there is a risk that the Strategy may not change how agencies approach evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, or that only some agencies may follow the guidance. As such, a monitoring process that assesses the extent that agencies are adhering to the Strategy is important. As Jobs Australia said:

Compliance with the Strategy is central to its success. (sub. 57, p. 10)

External oversight provides an incentive for agencies to ensure that the evaluations they undertake are high quality, are useful and that their evaluation processes, practices and outputs align with the Strategy's principles. This is a change from current arrangements.

However, monitoring agencies' performance against the Strategy should be much more than a compliance-driven exercise that mechanically assesses whether each and every agency has adhered to each aspect of the Strategy. In fact, the Strategy will effectively fail if it only results in agencies 'ticking the boxes' and viewing the Strategy as a compliance 'hurdle' that needs to be overcome. Instead, the focus should be on the outcomes that have been achieved because agencies have effectively complied with the Strategy.

Monitoring arrangements should also be a platform where the use of good evaluation practices by agencies are acknowledged, applauded and encouraged, and where lessons about what does (or does not) work well when evaluating policies and programs affecting Aboriginal and Torres Strait Islander people are shared. They should also be an avenue for sharing innovations in evaluation or engagement practices or approaches (including in cases

where innovations have not worked, but good practice was followed). The benefit of such a function was pointed out by the Lowitja Institute:

Given limited formal reporting of what has and hasn't worked in terms of approaches, the allocation of time and resources to identify and record the lessons learnt from various evaluations approaches would be of significant value. (sub. 50, p. 8)

Sharing lessons (and the integration of these into agencies' practices) will be important for creating a culture where evaluation results are valued and for deepening the capability of agencies to effectively evaluate policies and programs affecting Aboriginal and Torres Strait Islander people.

Monitoring the effectiveness of the Strategy itself

The objective of the Strategy is to improve the lives of Aboriginal and Torres Strait Islander people by having policy and program decisions informed by high-quality and useful evaluation evidence. It is important that monitoring arrangements assess whether the Strategy itself encourages high-quality and useful evaluations in pursuit of that aim.

The Strategy (and the accompanying *A Guide to Evaluation under the Indigenous Evaluation Strategy*) will be most effective if it is regularly updated to ensure it continues to reflect what is considered to be good practice evaluation (knowledge on what is good practice should deepen over time as agencies progress with implementing the Strategy and better evaluation is undertaken).

And because the Strategy is new and untested, it may become apparent that some parts of the Strategy work well and some parts less so as agencies begin to implement it. Regular monitoring should identify where the Strategy is not meeting expectations (and why) and how the Strategy's implementation can be improved to ensure it continues to be fit-for-purpose and useful.

Some principles to underpin the design of monitoring arrangements

Monitoring arrangements should be independent ...

Enshrining independence in the monitoring arrangements for the Strategy is critical if they are to be viewed with credibility. Independence means that those doing the monitoring — free from the day-to-day operation of government or service delivery — have greater scope to candidly examine all aspects of the Strategy's implementation (both good and bad). Independence also means that those monitoring the Strategy do not have opportunity to 'mark their own homework' by passing judgement on their own performance in implementing the Strategy, and as such, that outputs from the monitoring process are more likely to be viewed as impartial and valid.

... consultative ...

A lot of the knowledge needed to inform the monitoring of the Strategy will reside with those who are involved in its implementation. As such, it is important that monitoring arrangements provide opportunities for people to share this knowledge.

Views on the performance of the Strategy should be sought from a wide range of people, including Aboriginal and Torres Strait Islander people, government agencies, program managers and evaluators.

... transparent ...

The information feeding into (to the greatest extent possible) — and the outputs that stem from — the monitoring process should be public. Transparency provides a degree of accountability on those doing the monitoring to ensure that they are rigorous and effective in fulfilling this role.

Transparent monitoring arrangements also provide additional incentives for agencies to ensure they are effective in implementing the Strategy (else risk being called out publicly). As Maggie Walter, commenting on the Australian National Audit Office's (ANAO's) report on the Indigenous Advancement Strategy, said:

No agency wants a report written about them that says they are not doing a good job.
(sub. 112, p. 1)

... and sufficiently resourced

Monitoring arrangements should have access to the resources needed to be comprehensive, rigorous and consultative. This goes beyond sufficient funding — although this is important — but also includes sufficient time and access to the mix of skills needed to monitor the Strategy effectively. It also means those monitoring the Strategy have sufficient authority to undertake the role.

Proposed monitoring arrangements for the Strategy

Who should monitor the Strategy?

Reflecting the principles outlined above, for monitoring arrangements to be most effective, those tasked with monitoring the Strategy ideally would:

- not be required to undertake evaluations under the Strategy, otherwise the monitoring process will be susceptible to criticism that it allows a degree of self-assessment and will not be viewed as credible

-
- be independent of the day-to-day operation of government, so it can candidly examine the performance of government agencies against the Strategy. Ideally, this would take the form of a statutory authority or office, underpinned by legislation so that their independence and authority is formally established and enshrined. This would provide clarity about the authority's remit and powers, and provide them with a legislative basis to undertake their work with confidence and without fear of reprisal. Because legislation can only be changed by the Parliament, it is also harder for a statutory authority to be unwound or its remit altered, meaning they would have more certainty and scope to make longer-term investments to carry out the monitoring role more effectively
 - have Indigenous governance, consistent with the overarching principle of the Strategy of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges
 - have expertise and experience in evaluation, and in particular, knowledge of what constitutes good evaluation in the context of policies and programs affecting Aboriginal and Torres Strait Islander people
 - be respected as a 'honest broker'. They should be viewed as impartial by those with an interest in the Strategy, including Aboriginal and Torres Strait Islander people, the Australian Government, government agencies, and the community more broadly. If this is not the case, there is a risk that the conclusions and recommendations from monitoring arrangements are dismissed as biased (and ignored), and that they do little to incentivise agencies to implement the Strategy effectively
 - be consultative and transparent in the way they conduct their business.

As discussed in chapter 7, while the Letter of Direction implies that the Commission would have a role reviewing the performance of agencies against the Strategy, we are proposing a new unit — we have called it the Office of Indigenous Policy Evaluation (OIPE) — be established to oversee the Strategy's implementation. This oversight role should include monitoring agencies' performance against the Strategy and the effectiveness of the Strategy itself. It is proposed that the OIPE initially be housed within an existing independent statutory authority.

Housing responsibility for monitoring the Strategy's implementation in an independent statutory authority is important if monitoring is to be viewed as impartial and credible. But there are risks and challenges with housing monitoring arrangements in a pre-existing authority whose established remit, powers, capabilities and processes may not be well-aligned to the monitoring task.

The Commission is proposing — in the longer term — a new statutory authority (the Centre for Evaluation Excellence or CEE) be established to oversee and support quality evaluation of social and health policy across the Australian Public Service (APS) (chapter 7), and that the OIPE should be transferred to the CEE once it is established. This would, in effect, place responsibility for monitoring the performance of the Strategy, and the conduct of agencies

against it, within the remit of the CEE. This has several potential advantages, vis-à-vis retaining monitoring arrangements for the Strategy in an existing statutory authority.

- First, as part of establishing the CEE, the Australian Government has the opportunity to build the monitoring role for the Strategy into the CEE's design (including in the design of the legislation that underpins the CEE). This means that there is greater scope to tailor and formalise the structure, authority and powers of the CEE so that they have all the tools they require to exercise the monitoring role effectively.
- Second, given the primary remit — and the primary accountability — of the CEE would be to oversee and improve evaluation practice in social and health policy across the APS, the CEE would have greater scope (and greater incentives) to gather and build evaluation expertise within their agency, including the expertise needed to undertake the monitoring role effectively.
- Third, as discussed in chapter 7, there are benefits to extending oversight and monitoring arrangements for evaluation to other policy areas beyond just evaluations that affect Aboriginal and Torres Strait Islander people. The establishment of the CEE provides a clear opportunity to expand monitoring arrangements for other social and health policy evaluations undertaken across the APS.

A proposed approach

Irrespective of where the OIPE sits, the Office — in partnership with the Indigenous Evaluation Council (chapter 7) — should produce regular reports on the state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. These reports should include findings and recommendations based on their monitoring of the Strategy's implementation. The reports should be publicly released.

Transparency is an important aspect of the monitoring process and for raising the bar on the quality and usefulness of evaluations. The potential for agencies' good processes and practices (including, for example, comments from evaluation participants) to be highlighted publicly provides an additional incentive for good practice evaluation. This transparency is also important for the credibility of the monitoring process, because it allows for the performance of those doing the monitoring to be publicly scrutinised.

How should agencies' performance be assessed against the Strategy?

The Commission is proposing that agencies' performance against the Strategy be assessed in two main areas:

- the quality of agencies' rolling Three Year Evaluation Forward Work Plans (chapter 6)
- the quality and usefulness of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people undertaken by agencies, including whether completed evaluations have been publicly released.

An indicative approach, including actions and activities that could be undertaken to make these assessments, and what the assessments might look like, is set out in table 10.1 (for rolling Three Year Evaluation Forward Work Plans), and table 10.2 (for completed evaluations).

A maturity approach should be adopted for these assessments, recognising that agencies will be starting from different bases, and for some, it will take time to develop good practice. A maturity approach — in the context of evaluation and engagement — is outlined in the draft Strategy. The practices and cultures associated with each level of maturity will be developed in more detail following the release of the draft Strategy, with the aim of providing agencies with a clear idea of what success looks like and a standard against which agencies' progress in implementing the Strategy can be measured.

Agencies who have not upheld their basic responsibilities under the Strategy should be identified, such as agencies who have not prepared and/or published a rolling Three Year Evaluation Forward Work Plan, or those who have not published completed evaluation reports (unless there is a valid reason not to).

As noted in chapter 6, the monitoring process is also an opportunity to establish a formal set of evaluation priorities to help guide where agencies should focus their evaluative efforts and resources in the future (table 10.3). In the first instance, the 'interim' evaluation priority areas should be built on, and an assessment undertaken as to whether they are still appropriate. Formal evaluation priorities — including opportunities for cross-agency evaluations — should continue to be better defined over time as better information becomes available, including information on:

- the policy and reform priorities agreed to by the Joint Council on Closing the Gap
- the evaluation and policy landscape of agencies as they report better quality information through rolling Three Year Evaluation Forward Work Plans and completed evaluation reports on key policies and programs
- other evaluation priorities identified through consultations as part of this monitoring process, and with the benefit that this process is undertaken as a joint partnership governance arrangement with Aboriginal and Torres Strait Islander people.

Table 10.1 Assessing the quality of rolling Three Year Evaluation Forward Work Plans — an indicative approach

<i>Task</i>	<i>Activities for the monitoring authority</i>	<i>Outputs</i>
Assessing whether agencies have published a rolling Three Year Evaluation Forward Work Plan.	• Examining agency websites to determine whether they have published a rolling Three Year Evaluation Forward Work Plan. • Requesting information from agencies about whether they have published a plan.	• Reporting on the level of compliance with the requirement to publish a rolling Three Year Evaluation Forward Work Plan. • Reporting on agencies that have not met the requirement to publish a plan.
Assessing the extent that rolling Three Year Evaluation Forward Work Plans are high quality, including the extent that: • the methodology for how agencies have categorised priority programs is clear and reasonable • agencies have developed and used processes to reflect Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in their plans • plans are clear on what policies and programs agencies intend to evaluate and when and how they will be evaluated (or how agencies intend to prepare programs for evaluation) • planned evaluations map to priority areas • plans contain an acceptable mix of new and existing programs to be evaluated.	• Assessing the quality of all rolling Three Year Evaluation Forward Work Plans. • Focussing on a sample of plans for a more detailed assessment of their quality. • Seeking information from agencies about the processes they have undertaken to develop their plans. • Seeking views from Aboriginal and Torres Strait Islander people (and/or the community more broadly) on how plans have been developed.	• Identification of rolling Three Year Evaluation Forward Work Plans that represent good practice and/or case studies detailing how good practice can be undertaken. • Reporting on instances where plans have resulted in a more systematic approach to setting evaluation priorities, and channelled evaluation resources and efforts to high priority areas. • Recommendations for areas of improvement for future plans.

Table 10.2 Assessing the quality and usefulness of completed evaluations — an indicative approach

<i>Task</i>	<i>Activities for the monitoring authority</i>	<i>Outputs</i>
Assessing the extent that evaluations undertaken match those identified in rolling Three Year Evaluation Forward Work Plans and high priority evaluations are taking place.	• Comparing agencies' completed evaluations with those identified in rolling Three Year Evaluation Forward Work Plans, and identify differences. • Consulting with agencies about why particular planned evaluations were not undertaken.	• Reporting on the extent that agencies are following their rolling Three Year Evaluation Forward Work Plans. • Identification of particular agencies where plans do not appear to influence evaluations undertaken.
Assessing the extent that evaluations undertaken under the Strategy are high quality, including the extent that: • the design, conduct and management of evaluations is undertaken in line with the principles of the Strategy. This includes that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges have been centred in evaluations (including in commissioning arrangements) • there are capability gaps that appear to be impeding evaluation quality.	• Undertaking a high-level analysis of the quality of all evaluations undertaken under the Strategy. • Focussing on a sample of evaluations for a more detailed assessment of their quality. • Assessing the statements made by agencies on how they have adhered to the principles of the Strategy in their evaluations. • Seeking information from agencies on how they commissioned and/or undertook evaluations. • Seeking views from Aboriginal and Torres Strait Islander people (and the community more broadly) on the quality of evaluations.	• Reporting on the overall quality of evaluations undertaken under the Strategy. • Identification and highlighting of evaluations that have adopted innovative approaches and/or represent good practice, and/or case studies detailing how good practice can be undertaken. • Reporting on instances where evaluations have resulted in improvements to policies and programs affecting Aboriginal and Torres Strait Islander people. • Recommendations on how agencies should commission or undertake future evaluations. • Recommendations on how agencies should address capability gaps in evaluation.
Assessing the extent that agencies are publishing their evaluation reports, including the extent that: • published evaluations include a short plain English summary • evaluation reports and summaries document how the planning, commissioning and conduct of the evaluation adhered to the principles of the Strategy • published evaluations include a management response and action plan.	• Assessing what proportion of evaluations have been published in the Indigenous Evaluation Clearinghouse. • Assessing what proportion of published evaluations document how the Strategy's principles have been adhered to, and what proportion contain management responses and action plans. • Requesting information from agencies about evaluations they have or have not published.	• Reporting on the publication of evaluation reports. • Identification of agencies that have not met the requirement to publish evaluations. • Recommendations about how future evaluations should be published.

Table 10.3 Formalising evaluation priorities — an indicative approach

<i>Task</i>	<i>Activities for the monitoring authority</i>	<i>Outputs</i>
Refining evaluation priorities to guide agencies' evaluation resources and efforts.	• Examining whether existing (and in the first instance, 'interim') evaluation priorities are appropriate, guided by: - improved evidence on policy impacts from better quality evaluations - new governance arrangements to have Aboriginal and Torres Strait Islander people as partners in the process of identifying priorities - consultation on priorities with agencies and key stakeholders - improved information and better coordination of existing evaluation activity through published rolling Three Year Evaluation Forward Work Plans.	• Recommendations to government on a formal set of government-wide evaluation priorities.
Identifying where there may be opportunities for collaboration on evaluation across agencies.	• Examining the content of rolling Three Year Evaluation Forward Work Plans to identify areas of commonality in plans across agencies. • In instances where common evaluation areas are identified, assessing whether there may be benefits from agencies collaborating in these evaluations (such as improved evaluation coverage, greater scope to examine interactions between policies and programs and/or resource savings).	• Recommendations outlining areas where collaboration should occur.

How should the Strategy's effectiveness be monitored?

Monitoring the Strategy's effectiveness should be about examining the extent to which the Strategy has encouraged high-quality and useful evaluations. This will involve assessing whether the Strategy (and the accompanying *A Guide to Evaluation under the Indigenous Evaluation Strategy*) has encouraged agencies to undertake more effective evaluation. It should also include looking at whether there is scope to improve the Strategy's implementation and/or to update or enhance the content of the Strategy or its accompanying guidance material. This is consistent with the need for the Strategy to be a living document that is updated to reflect changes in what is considered to be good practice, and to expediently address any issues in the way that the Strategy is working that threaten to reduce its usefulness and value (this is especially important because the Strategy will be new and untested).

Monitoring the Strategy's effectiveness will involve considering the extent that:

- the Strategy's priority setting process is directing evaluation effort to high priority areas
- the way agencies are implementing the Strategy is resulting in Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges being centred in evaluations affecting Aboriginal and Torres Strait Islander people
- the Strategy and its guidance material reflects the most up-to-date knowledge on good practice evaluation for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people
- the findings from evaluations undertaken under the Strategy are providing the information needed to influence the way policies and programs affecting Aboriginal and Torres Strait Islander people are designed, implemented and reviewed.

Evidence generated from monitoring agencies' performance against the Strategy will provide some of the information needed to answer these questions. But insights from the experiences of those involved with the Strategy will also be important for reaching conclusions on the Strategy's effectiveness. This includes the experiences of those overseeing, doing, or participating in evaluation, as well as those impacted by evaluations undertaken under the Strategy.

For this reason, it is important that the approach to monitoring the Strategy's performance includes avenues for people to share their views on how the Strategy is working. A public submission process should be one such avenue. In addition to accepting submissions, targeted consultation, roundtables and workshops should also be considered.

Recommendations could also be made if opportunities are identified to improve how the Strategy is implemented. These recommendations could be directed at agencies, or — in the event that changes at a whole-of-government level could facilitate better implementation — the Australian Government.

When and how often should monitoring reports be produced?

Answering these questions highlights several trade-offs.

On the question of *when the first monitoring report should be publicly released*, the more time there is between when the Strategy is endorsed and when the first report is released, the more information there will be to assess agencies' performance against the Strategy (as well as to assess the performance of the Strategy itself).

However, if the first monitoring report is too far into the future, agencies may lack the incentive to change how they evaluate policies and programs affecting Aboriginal and Torres Strait Islander people (in effect, 'putting it off until later'). The timing of the first monitoring report for the Strategy needs to balance these two somewhat competing considerations.

On the question of *how often monitoring reports should be produced*, more regular reporting provides more opportunities to highlight agencies' performance against the Strategy, and to update the Strategy to ensure it reflects good practice.

However, if reporting is too frequent, there is a risk that monitoring can only occur at a fairly superficial level. This would be far from ideal, given much of the value from the proposed monitoring process comes from: 'getting into the weeds' about the quality of evaluations; identifying lessons and good practice; and examining what is and is not working with the Strategy itself.

More frequent reporting also runs the risk of 'consultation fatigue.' Effective consultation is a prerequisite for high-quality and credible monitoring arrangements. But effective consultation relies on the goodwill and resources of those being consulted. If reporting is done too regularly, and demand for consultation is too high, participants may not have the resources to contribute their views and experiences to inform the monitoring of the Strategy.

Recognising these trade-offs, we propose that the first monitoring report be provided to the Australian Government (and released publicly) two years after the Strategy is first endorsed.

A report two years after the Strategy is endorsed should be enough time for agencies to start reflecting the principles and guidance of the Strategy into their evaluation processes, practices and products. While two years' time may be too soon to examine (in detail) the quality of evaluations undertaken under the Strategy (there may be too few completed evaluations to allow for informed judgements), there should be more than enough information to make assessments on the performance of agencies in completing their rolling Three-Year Evaluation Forward Work Plans, as departments and major agencies will be expected to have an initial Plan within a year of the Strategy being implemented (chapter 6).

After the first report is released, monitoring reports should be produced biennially (draft Strategy, action 11). While the first report is expected to focus on evaluation planning, subsequent reports can be expected to have a greater emphasis on assessing the quality of

evaluations undertaken (or commissioned) by agencies, including the extent that the Strategy is effective in promoting high-quality and useful evaluations.

10.2 Review arrangements for the Strategy

What do review arrangements for the Strategy need to achieve?

As discussed above, the Commission is proposing that the OIPE — sitting within a statutory authority — be responsible for the ongoing monitoring of the Strategy’s implementation and that it should produce biennial reports that assess both agencies’ performance against the Strategy and the performance of the Strategy itself. These reports may also include recommendations to improve how the Strategy is implemented or to update the Strategy to ensure it represents good practice.

However, along with the ongoing monitoring process, it is also important that a separate process for reviewing the Strategy be established. This review process should assess:

- whether the Strategy’s principles remain fit for purpose
- the extent that the Strategy has been effective in encouraging higher quality and more useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, including the effectiveness of the Strategy’s actions
- the extent that the performance of those tasked with overseeing the Strategy has been effective and whether changes to the policy environment require changes to oversight arrangements.

In effect, the review process is an opportunity to examine whether the fundamental elements of the Strategy — the Strategy’s principles, actions, and oversight arrangements — are effective and whether improvements can be made to these. This means that the review process involves broader and deeper assessments and analysis of the Strategy’s effectiveness than what will be undertaken through the monitoring process, and the level of resourcing for the review process should support this.

Assessing whether the principles of the Strategy remain fit-for-purpose

The principles are the key building blocks of the Strategy. They set out what high-quality practice looks like and what agencies should consider when undertaking evaluations, while also allowing flexibility for evaluations to be tailored to the particular circumstances of policies, programs and communities. Given that the principles seek to influence all aspects of evaluation — including priority setting, evaluation planning, evaluation conduct, commissioning practices and evaluation reporting — it is very important that they remain fit-for-purpose and useful.

The monitoring process is well placed to assess whether the principles, as set out in the Strategy, are being upheld (or whether more guidance on their application is required). However, it is less well placed to assess whether the principles remain the right ones to deliver evaluations that provide high-quality and relevant evidence that can be drawn upon to improve the lives of Aboriginal and Torres Strait Islander people. There is also benefit in having a review process that examines the efficacy of the principles that is independent from the monitoring process.

Assessing the extent the Strategy has been effective in encouraging high-quality and useful evaluations, including the effectiveness of the Strategy's actions

A review process also provides an opportunity to consider the extent that the Strategy has been effective in encouraging high-quality evaluations, and to provide recommendations to the Australian Government if opportunities to improve the Strategy are identified. The findings from the monitoring process will provide useful inputs to help inform such an assessment, although evidence will need to be collected from a range of other sources, including Aboriginal and Torres Strait Islander people and government agencies who are implementing and using the Strategy.

As part of assessing the Strategy's effectiveness, the review should consider the effectiveness of the actions required by agencies as set out in the Strategy. These actions cover steps that agencies should take to implement the Strategy effectively — including: developing a process to determine evaluation priorities; publishing rolling Three Year Evaluation Forward Work Plans; undertaking Indigenous Evaluation threshold assessments for new policies and programs that are assessed as having a significant and/or differential impact on Aboriginal and Torres Strait Islander people; and publishing evaluation reports with a management response and action plan.

The monitoring process will assess agencies performance in implementing these actions, and provide recommendations on how their implementation can be improved. But the monitoring process is less well-placed to assess the more fundamental question of whether each action remains fit-for-purpose. For example, the monitoring process might be able to identify whether changes to the content of rolling Three Year Evaluation Forward Work Plans would be beneficial, but it is less well placed to assess whether rolling Three Year Evaluation Forward Work Plans remain the most effective way to encourage agencies to take a transparent and systematic approach to their evaluation planning.

Assessing the extent that oversight arrangements remain fit-for purpose

As discussed above, the Commission is proposing a new unit — the OIPE — be established to oversee the Strategy. This oversight role includes:

- championing the value of evaluation in policy making and identifying potential areas for evaluations or systematic reviews on cross-cutting topics

-
- providing advice to Australian Government agencies on what the Strategy means in practice
 - identifying evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people
 - overseeing the development and progress of agencies' rolling Three-Year Evaluation Forward Work Plans
 - monitoring and publicly reporting on Australian Government agencies' performance against the Strategy and on how effective the Strategy has been in improving the quality and usefulness of evaluations affecting Aboriginal and Torres Strait Islander people (chapter 7).

The OIPE would exercise their functions with strategic guidance from, or in partnership with, the Indigenous Evaluation Council.

Regardless of where these oversight functions are housed, it is essential that they work effectively. While the Strategy ultimately places the onus on agencies to undertake better evaluation of policies and programs that affect Aboriginal and Torres Strait Islander people, these oversight arrangements help to provide the tools and the incentives agencies need to implement their responsibilities effectively.

A review process is important because it provides an opportunity to examine how effective the oversight arrangements have been and whether there is scope to improve them. A review process should consider the following questions.

- *Has the performance of those tasked with overseeing the Strategy been effective?* Assessing the performance of those overseeing the Strategy is valuable for two reasons. First, knowing that their performance will be subject to external review provides an incentive for those overseeing the Strategy to invest the resources needed to undertake the oversight role well. Second, external review is an opportunity to identify improvements in the way those overseeing the Strategy go about their work and that could potentially make oversight arrangements more effective.
- *Has there been changes to the policy environment that require changes to the oversight arrangements for the Strategy?* As identified in chapters 1 and 2, the policy landscape in which the Strategy will operate is dynamic and subject to considerable change. Initiatives are underway that have the potential to impact on the Strategy including:
 - the *Partnership Agreement on Closing the Gap*, including the commencement of three-yearly evaluations of Closing the Gap as agreed to by the Coalition of Aboriginal and Torres Strait Islander Peak Organisations and the Council of Australian Governments (JCOCTG 2019c, p. 7)
 - actions being undertaken by the Australian Government as part of its response to the Independent Review of the APS (although these actions are currently paused as a result of the COVID-19 pandemic (Gaetjens and Woolcott 2020, p. 2)). These actions include establishing a small team in the Department of Finance to help build

evaluation expertise and practices, and establishing an evaluation profession within the APS (chapter 2).

A review process is an opportunity to consider how the outcomes of these initiatives — and/or changes to how evaluation is undertaken in the APS more broadly — affect how oversight arrangements for the Strategy are best structured.

When should the Strategy be reviewed?

For the review of the Strategy to be most effective, there needs to be enough information on how it, and the oversight arrangements for it, are working. This suggests a longer period between when the Strategy commences and when it is first reviewed.

However, if too much time elapses before the Strategy is reviewed there is a risk that it may become ‘out of date’, given the policy environment around the Strategy is changing significantly. It is important that the review process is able to consider how the Strategy best interfaces with these changes.

Given this balance, the Commission is proposing that the Strategy should be reviewed five years after it is first endorsed by the Australian Government (draft Strategy, action 12). This would mean that the review of the Strategy would commence after the second monitoring report has been produced. Five years should be long enough to allow for valid judgments about the effectiveness of the Strategy and how oversight arrangements are operating, but soon enough to ensure that the Strategy continues to be valuable and useful in a changing policy landscape.

Some important features of the review process

Given we are proposing the review process be undertaken five years after the Strategy is endorsed, we have not been overly prescriptive about who should undertake the review or about how the review should be designed. These are decisions best made by the Australian Government closer to the review’s commencement. That said, there are several features that the Commission considers are important if the review is to be done well. These features relate to:

- *independence* — those undertaking the review of the Strategy should not be involved in its implementation. This independence is essential if the review process is to be viewed with credibility
- *transparency* — the review’s findings and recommendations should be made public
- *consultation* — the review should consult widely and seek public submissions. It should seek views from Aboriginal and Torres Strait Islander people, Australian Government agencies and people with views on how the Strategy may be improved. Those undertaking the review should also consult with those overseeing the Strategy to get their views on how oversight arrangements are working.

11 Bringing it all together: an Indigenous Evaluation Strategy

Key points

- The Indigenous Evaluation Strategy should include the following key components:
 - a clearly defined objective
 - a principles-based framework for evaluation, with practical guidance on applying the principles across all stages of evaluation
 - an outline of key roles and responsibilities, including Aboriginal and Torres Strait Islander governance arrangements
 - a process for determining evaluation priorities, both across the Australian Government and within agencies
 - actions to support a culture of evaluation
 - a process for monitoring and reviewing progress in implementing the Strategy.
- The overarching principle of the Indigenous Evaluation Strategy should be centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. The overarching principle is the lens through which the Strategy's other proposed principles — credible, useful, ethical and transparent — are interpreted.
- The Strategy should include a maturity approach to implementing the principles, which provides agencies with a roadmap for improving their evaluation practices.
- The Australian Government should endorse the Strategy and mandate that Australian Government agencies comply with it.

This chapter brings together insights from earlier chapters of this background paper and looks at what should be covered in the Indigenous Evaluation Strategy (the Strategy).

As discussed in chapter 1, the Strategy has not been co-designed with Aboriginal and Torres Strait Islander people as some participants suggested it should be. However, the Commission has drawn extensively on the views of participants, including those expressed in meetings, submissions and roundtables.

This chapter first outlines the key components of the Strategy (section 11.1) and then discusses the proposed principles for the Strategy (section 11.2).

11.1 The key components of the Strategy

The main purpose of the Strategy is to provide direction and guidance to Australian Government agencies on how to select, plan, conduct and use evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. However, the Strategy should also seek to help agencies build evaluation capability and a culture of evaluation (that is, a culture that embeds evaluative thinking within an agency, chapter 8).

As the Australian National Audit Office's review of the Indigenous Advancement Strategy (IAS) Evaluation Framework pointed out, a range of supports are needed to improve evaluation culture:

Good evaluation practice includes ... Embedding a comprehensive approach to evaluation practice, such as training, guidance, follow up and governance/oversight arrangements, to enable an evaluation culture to be established. (ANAO 2019b, p. 11)

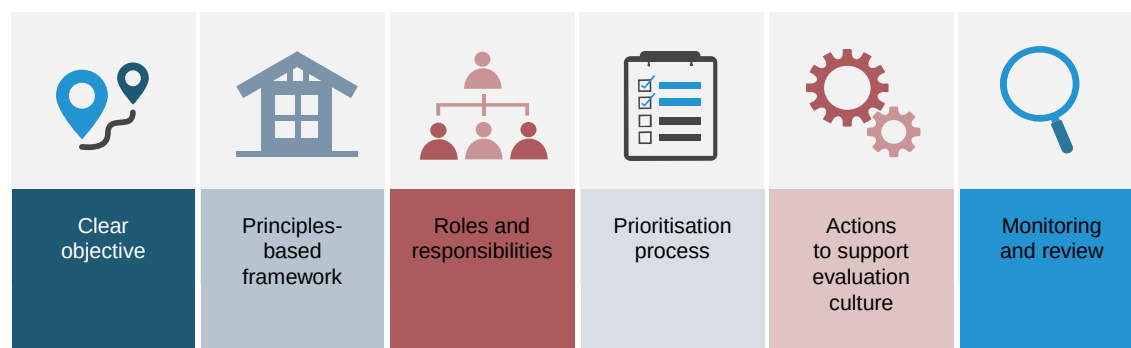
The Letter of Direction asks the Commission to establish a principles-based framework, identify priorities for evaluation, and set out an approach to reviewing agencies conduct of evaluations against the Strategy. While these are three key components of the Strategy, on their own they are unlikely to bring about the changes required to evaluation practices and culture across Australian Government agencies that will make the Strategy effective. The Strategy also needs a clearly defined objective, 'actions' to support a culture of evaluation across the Australian Public Service and a clear articulation of who is responsible for what under the Strategy.

In sum, the key components of the Strategy should include:

- a clearly defined objective
- a principles-based framework for evaluation, with practical guidance on applying the principles across all stages of evaluation
- an outline of key roles and responsibilities, including Aboriginal and Torres Strait Islander governance arrangements
- a process for determining evaluation priorities, both across the Australian Government and within agencies
- actions to support an evaluation culture
- a process for monitoring and reviewing progress in implementing the Strategy (figure 11.1).

Each of these components is discussed in more detail below.

Figure 11.1 Key components of the Indigenous Evaluation Strategy



A clearly defined objective

A strategy should have a clearly stated objective, or put another way, identify what success will look like. As discussed in chapter 1, there is strong agreement that the objective of the Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people. The Strategy should also make clear how higher quality and more useful evaluations will improve the lives of Aboriginal and Torres Strait Islander people.

A principles-based framework

The use of principles to guide evaluation practice is commonplace. Professional evaluation associations typically provide principles-based guidance on ethical conduct and professional standards. Examples include the Australasian Evaluation Society's *Guidelines for the Ethical Conduct of Evaluations* (AES 2013b) and the American Evaluation Association's *Guiding Principles* (AEA 2018).⁴¹ Principles-based frameworks for evaluation are also used in other Australian and state and territory government evaluation strategies (chapter 2).

There was widespread support among participants for a principles-based framework to guide evaluation practices under the Strategy. Participants saw the advantages of a principles-based approach as providing a shared view of what good practice evaluation looks like, while allowing flexibility for evaluations to be tailored to the particular circumstances of policies, programs or communities. For example, Chelsea Bond et al. said:

The power of principles for policy development, program delivery and evaluation is that they are adaptable to different contexts. While predefined structures and contracted outcomes appear tractable to high-level management, they are not meaningful or effective at the local level. Principles allow both the local-level experts and high-level abstract managers to have a shared

⁴¹ The Australasian Evaluation Society is now known as the Australian Evaluation Society.

understanding of how ‘good’ can be judged. Thus, principles can be framed in a way that can guide decisions, operations and evaluation across all levels ... (sub. 40, p. 11)

The Smith Family also said:

The Framework must be broad enough to incorporate the very diverse contexts in which programs and policies impacting on Aboriginal and Torres Strait Islander people occur.

A principles-based framework allows both for:

- The flexible application of different approaches and methods as appropriate to local contexts, and
- An explicit statement of the values and principles that underpin evaluations, which will guide their conduct. (sub. 56, p. 5)

And Empowered Communities argued that a principles-based approach would allow the Indigenous Evaluation Strategy to adjust to a changing policy landscape:

Given the dynamic nature of the reform process being undertaken in many of the place-based, collective impact and co-design models that government is currently funding, the use of principles-based evaluation models would serve to embed methodologies that are able to manage the dynamic nature of the policy frame. (sub. 41, p. 11)

While recognising the value of principles to guide evaluation practice, a number of participants also highlighted the need for practical guidance on how to implement the principles (box 11.1).

Box 11.1 What participants said about the need for practical guidance on implementing the Strategy’s principles

The Department of Health:

There is an opportunity for the Strategy to offer practical guidance in terms of principles, approaches and methodologies that respect the perspectives of Aboriginal and Torres Strait Islander people in the design, delivery and evaluation of Australian Government policies and programs. (sub. 35, p. 1)

The Western Australian Government:

Practical resources may also need to be made available to support Government agencies to translate the evaluation framework into practice. (sub. 74, p. 2)

Maggie Walter:

A principles based approach needs to conceptualise those principles along with examples of how they can be operationalised. They need to make clear the meaning of the principle and how can it be implemented. (sub. 112, p. 3)

The Commission agrees that principles, without guidance on what they mean in practice, will be of little value to Australian Government agencies, and are unlikely to be effective at raising the bar on the quality of evaluations or improving their usefulness. As such, the Strategy provides practical guidance to agencies on how the principles should be applied at all stages of evaluation, as well as ways to build capability and a culture of evaluation (figure 11.2). Balancing the need for the Strategy to provide practical guidance and a desire for it to be a

concise document, the Commission has also produced a companion document — *A Guide to Evaluation under the Indigenous Evaluation Strategy* — to provide further guidance to agencies on implementing the principles, over and above that provided in the Strategy.

Figure 11.2 Practical guidance is needed on applying the principles at every stage of evaluation



Clearly defined roles and responsibilities

Strategies usually also define roles and responsibilities. This is because a document that sets out objectives and principles will not be effective without governance arrangements that clearly define who is responsible for what and that embed incentives to change behaviour.

While primary responsibility for conducting evaluations under the Strategy should continue to lie with Australian Government agencies, the Commission considers that centralised leadership, support, coordination and oversight is needed if the Strategy is to make a difference to evaluation processes and practice (section 7.3). The Strategy should outline the Commission’s proposed governance arrangements for the Strategy, including how such arrangements ensure effective engagement with Aboriginal and Torres Strait Islander people. A number of participants pointed to the importance of having an Aboriginal and Torres Strait Islander governance mechanism for the Strategy. For example, the Aboriginal Health Council of South Australia said:

... we strongly support the implementation of an Aboriginal governance structure be included as a main component of the overall strategy. (sub. 61, p. 2)

And Indigenous Community Volunteers suggested:

... the role of First Nations Peoples be added into the essential elements of the Strategy. ... First Nations governance should be considered for the Strategy as a whole and for each evaluation undertaken. (sub. 88, p. 8)

A process for setting evaluation priorities

As discussed above, the Letter of Direction asks the Commission to identify evaluation priorities as part of the Strategy. Government-wide evaluation priorities will help focus agencies’ evaluative efforts under the Strategy and ensure that agency-level priority setting is guided by

what is valued most by Aboriginal and Torres Strait Islander people. A more systematic approach to identifying evaluation priorities at the agency-level is also required (chapter 6).

Ideally, government-wide evaluation priorities would be identified in partnership with Aboriginal and Torres Strait Islander people to ensure that priorities reflect what is most important to improving the lives of Aboriginal and Torres Strait Islander people. The Victorian Aboriginal Child Care Agency recommended:

That government agencies engage with Aboriginal organisations and community representatives to determine evaluation priorities. This will enable better processes to draw out those programs and issues which are of priority to Aboriginal organisations and communities. (sub. 26, p. 6)

The Commission proposes that the Strategy outline a process for setting priorities — both at the government-wide and agency level. The government-wide priority-setting process should be done in partnership with Aboriginal and Torres Strait Islander people, through the Strategy’s proposed Indigenous Evaluation Council (chapter 7). Because it will take some time for the governance arrangements to be put in place, the draft Strategy should identify some interim priorities areas that can be refined between the draft and final Strategy.

A broad monitoring and review function

The monitoring of agencies’ conduct of evaluations against the Strategy will be important for the success of the Strategy. As discussed in chapter 10, external oversight of agencies’ conduct of evaluations against the Strategy will provide an incentive for agencies to ensure that the evaluations they undertake are of high quality, are useful, and that their evaluation processes, practices and outputs are aligned with the Strategy’s principles.

It is envisaged that monitoring under the Strategy will be more than checking agencies compliance with the Strategy and it will also focus on highlighting good evaluation practices and sharing lessons on what works (or does not work) well when evaluating policies and programs affecting Aboriginal and Torres Strait Islander people. Social Ventures Australia spoke about an ‘evaluation of evaluations’:

An Indigenous Evaluation Strategy could consider mechanisms for improving the quality, rigour and appropriateness of methodologies used in the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander peoples, such as an ‘evaluation of evaluations’ to better determine best practice. (sub. 83, p. 9)

The Commission is proposing that monitoring should also include an assessment of whether the Strategy itself encourages high-quality and useful evaluations in pursuit of its objective to improve the lives of Aboriginal and Torres Strait Islander people. Regular monitoring should identify where the Strategy is not meeting expectations (and why) and how the Strategy’s implementation can be improved to ensure it continues to be fit-for-purpose and useful (chapter 10). This broader emphasis on monitoring how well the Strategy is

functioning was supported by participants. The Aboriginal Health Council of Western Australia, for example, said it:

... supports the Productivity Commission's position that the Indigenous Evaluation Strategy be subject to ongoing review to ensure that it remains fit for purpose. (sub. 42, p. 4)

And in addition to monitoring, the Commission is proposing that an independent review of the Strategy be undertaken five years after it is endorsed to assess whether the Strategy's principles remain fit-for-purpose, the performance of governance and monitoring arrangements, and the Strategy's success in encouraging higher quality and more useful evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people (chapter 10).

A commitment to monitoring and review processes should be embedded in the Strategy to emphasise their important role.

Actions to support an evaluation culture

The Strategy will be most effective in an environment where evaluation is valued for accountability, learning and evidence-based decision making, and where agencies have the capability and data to undertake high-quality evaluation. Chapters 7, 8 and 9 identify a range of mechanisms for supporting a culture of evaluation within Australian Government agencies. The Strategy should include these action items.

Australian Government endorsement

The importance of Australian Government endorsement of the Strategy was highlighted by several participants. Empowered Communities, for example, said:

Achieving the objective of a whole of government Indigenous Evaluation Strategy ... will require more than a high-quality Productivity Commission report. The government will need to act on recommendations for change and compel its departments and agencies to adopt and implement those reforms as quickly as possible. ... In terms of adoption of the Indigenous Evaluation Strategy across government, most fundamental to the success of this work-program, is buy-in from agencies and centralisation of an evaluation culture in design of policy and programs. ... If the mandate for policy and program evaluation is not strongly and uniformly embedded across agencies, the result is likely to be piecemeal and of varying usefulness in informing future policy and program design activities and of limited value to Indigenous people. (sub. 41, pp. 6–7)

And Children's Ground:

Without a clear plan and mandate for implementation of monitoring and evaluation we are likely to see little change in the current levels of evaluation, let alone change in the effectiveness of evaluation and First Nations leadership and empowerment in evaluation. (sub. 73, p. 4)

It is the Commission's view that the Strategy would be strengthened by the Australian Government endorsing it and mandating that Australian Government agencies comply with the Strategy.

11.2 What principles should underpin the Strategy?

How were the principles decided on?

The principles for the Strategy were developed based on:

- a review of what constitutes good quality evaluation conduct and use (chapters 4 to 9)
- existing evaluation frameworks used by government agencies (chapter 2)
- frameworks for Aboriginal and Torres Strait Islander research and evaluation (chapter 5)
- participants' views, including those expressed in submissions, meetings and at a roundtable on Objectives, Principles and Defining Success, held as part of the project.

There is a high degree of consistency in the principles included in existing Australian Government agency evaluation frameworks (figure 2.2). Principles that are common across multiple frameworks include that evaluation be integrated, fit-for-purpose, evidence-based, independent, transparent, timely and useful.

Few Australian Government agencies' evaluation frameworks include principles relating to ethical research conduct or engagement with Aboriginal and Torres Strait Islander people or other key stakeholders. The IAS Evaluation Framework (DPMC 2018b) is a notable exception.

The principles for the Strategy need to go beyond general evaluation good practice and guide agencies on how to design and conduct culturally safe evaluations that incorporate the perspectives and priorities of Aboriginal and Torres Strait Islander people. Frameworks for conducting research and evaluation with Aboriginal and Torres Strait Islander people (table 5.3) include the following principles: leadership, partnership and engagement; relationships, reciprocity and shared responsibility; respecting and valuing cultures and knowledges; equity, accountability and ethical practice; and evidence-based.

A number of participants argued for principles similar to those above to be incorporated into the Strategy (table 11.1).

Table 11.1 Principles for the Strategy suggested by participants

<i>Victorian Aboriginal Community Controlled Health Organisation</i>	
• Recognition • Self-determination • Community building: do-no-harm and strength-based approaches • Ethics	• Accountability • Transparency • Effectiveness • Cultural continuity
<i>Victorian Aboriginal Child Care Agency</i>	
• Self determination • Culturally informed approaches • Accountability • Strengths based approach and one that is centred on the aspirations of Aboriginal people	• Aboriginal-led collaboration and partnership • Person centred • Community engagement • Equity
<i>Aboriginal Health Council of Western Australia</i>	
• Empowerment and self-determination for Aboriginal people • Cultural security • Commitment to co-design • Commissioning for prevention and early intervention • Evaluation based on Aboriginal measures of success, not mainstream norms	• Aboriginal approaches to evaluation • Evaluation informing service commissioning and design • Evaluation without agendas • Feedback • Meaningful and responsible use of data • Compliance • A sustainable evaluation framework
<i>NSW Aboriginal Land Council</i>	
• Respectful • Responsive • Reciprocal • Competent	• Accountable • Transparent • Ethical • Integrity
<i>Australian Evaluation Society</i>	
• Place the emotional and social wellbeing of Indigenous people at the centre of policies and programs, and/or as a priority of key performance indicators • Strengths-based approach to evaluation • Indigenous empowerment and self-determination • Culturally safe evaluation practice	• Respect for diverse Indigenous perspectives • Ethical evaluation practice • Collaborative, inclusive evaluation practice that genuinely involves Indigenous people in policy, program design and evaluation • The outcomes of evaluations are communicated and reported to Indigenous people

Sources: Aboriginal Health Council of Western Australia (sub. 42, pp. 1–4); Australian Evaluation Society (sub. 49, pp. 22–23); NSW Aboriginal Land Council (sub. 91, pp. 6–7); Victorian Aboriginal Child Care Agency (sub. 26, pp. 2–3); Victorian Aboriginal Community Controlled Health Organisation (sub. 44, pp. 7–12).

Others advocated principles based on existing frameworks for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people to inform the Strategy, including:

- the Lowitja Institute’s *Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health* (Kelaher et al. 2018b), suggested by the National Aboriginal Community Controlled Health Organisation (sub. 95, p. 4), Indigenous Allied Health Australia

(sub. 31, p. 4), the Australian Evaluation Society (sub. 49, p. 23), the Aboriginal Health Council of South Australia (sub. 61, p. 6), Close the Gap Campaign Steering Committee (sub. 71, p. 4) and the Aboriginal Medical Services Alliance Northern Territory (sub. 81, p. 2)

- the BetterEvaluation *Ethical Protocol for Aboriginal and Torres Strait Islander Evaluation* (Gibb et al. 2019), suggested by the National Aboriginal Community Controlled Health Organisation (sub. 95, p. 4), the Aboriginal Health Council of South Australia (sub. 61, p. 6) and the Fred Hollows Foundation (sub. 14, p. 11).

Proposed principles for the Strategy

The principles for the Strategy have two purposes. To guide:

1. agencies and evaluators when they are planning, conducting, reporting and using evaluation
2. priority setting, knowledge sharing and translation, building evaluation capability, monitoring and review under the Strategy.

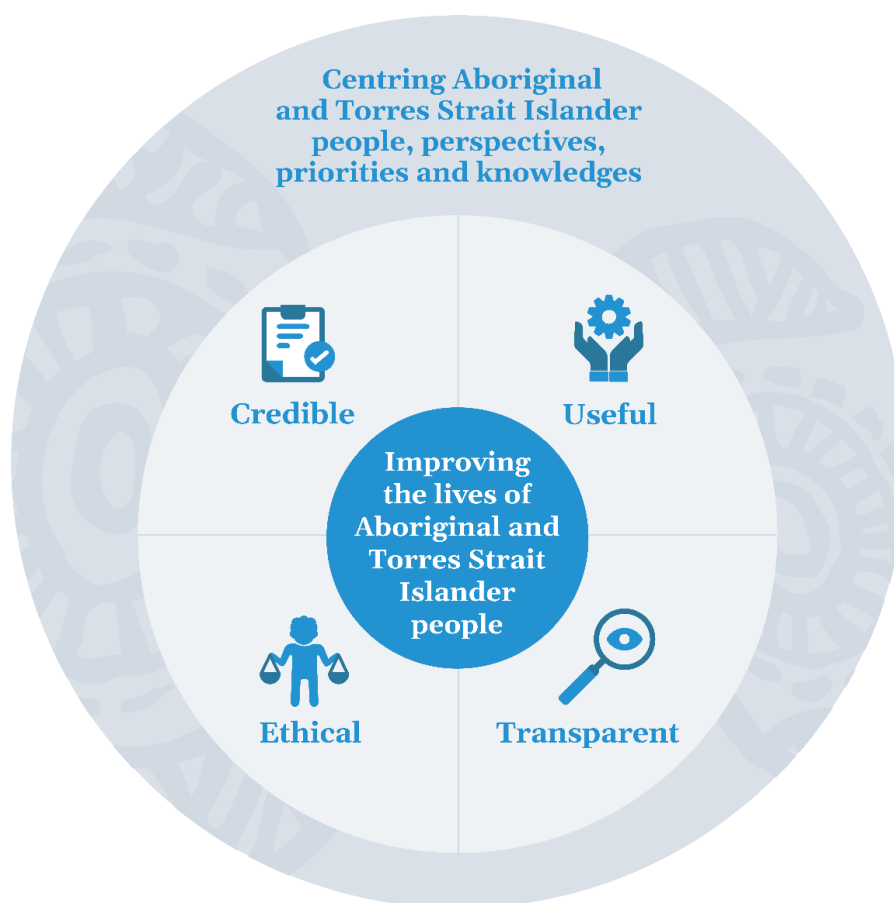
From the long list of possible principles, the Commission is proposing five principles aimed at lifting the bar on the quality of evaluations, enhancing their use and — importantly — putting Aboriginal and Torres Strait Islander people at the centre of evaluation and the Strategy. The advice of Empowered Communities was that:

The Commission should focus on agreeing a small number of principles — maximum 4-5 principles — that are focused, meaningful for all who will have to comply with them, and avoid the temptation to develop a long list of principles that encourage agencies and others to take a ‘tick-a-box’ approach. (sub. 41, p. 10)

Participants at the Objectives, Principles and Defining Success Roundtable were of a similar view about having a small number of principles for the Strategy.

The Commission proposes that the overarching principle of the Indigenous Evaluation Strategy be **centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges** and that the overarching principle is the lens through which the Strategy’s other proposed principles — **credible, useful, ethical** and **transparent** — are interpreted (figure 11.3).

Figure 11.3 **Guiding principles of the Indigenous Evaluation Strategy**



Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges

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

The proposed overarching principle is about recognising the strengths of Aboriginal and Torres Strait Islander people, communities, knowledges and cultures. It is about building genuine partnerships with Aboriginal and Torres Strait Islander people to define policy and program outcomes, decide on evaluation questions, how evaluations will be conducted and how evaluation findings will be interpreted.

Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges will improve the quality and use of evaluations of policies and programs and better align policies and programs with the needs and priorities of Aboriginal and Torres Strait Islander people.

The idea that Aboriginal and Torres Strait Islander people should be at the centre of the Strategy gives action to the calls by many participants for the principles to include: support for self-determination and empowerment; partnership, engagement and co-design; recognition of Aboriginal and Torres Strait Islander culture, knowledges and methodologies; and drawing on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people (table 5.3, table 11.1).

Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in the Strategy and in evaluation was a key recommendation of participants at the Commission's Principles, Objectives and Defining Success Roundtable. A number of submissions also called for Aboriginal and Torres Strait Islander people to be at the centre of the Strategy and evaluation.

In order to achieve effective evaluation and long term outcomes, we must place Aboriginal people and leadership at the centre of an [Indigenous Evaluation Strategy]. (Victorian Aboriginal Community Controlled Health Organisation, sub. 44, p. 11)

Ensuring First Nations people are at the centre of service and system design, delivery and evaluation and in progress reporting is possible, and is the only way to generate culturally appropriate and meaningful outcomes and measures for their children, families and community. When people are involved in and have ownership over services, programs and measurement they understand what data is being collected about them and engage in collecting and using the data to inform and drive locally relevant solutions. (Children's Ground, sub. 73, p. 10)

The most important objective for the Indigenous Evaluation Strategy should be for Aboriginal and Torres Strait Islander people to be at the centre of the process. To this end, the overarching principles of the Strategy should align with the United Nations Declaration on the Rights of Indigenous Peoples. (name withheld, sub. 3, p. 1)

As the key point of difference between the Strategy and most existing evaluation frameworks used by Australian Government agencies, evaluators and agencies are expected to require more guidance on implementing this principle than the other principles.

Credible

This principle is about ensuring that evaluation methods, processes and results are rigorous, fit-for-purpose, objective, independent and draw on the priorities, perspectives and knowledges of Aboriginal and Torres Strait Islander people. Credibility of methods, data and results is critical for ensuring that decision makers can rely on evaluation results to design, adjust and implement policies and programs affecting Aboriginal and Torres Strait Islander people.

Credible evaluations require evaluators who are technically and culturally capable, as well as independent of the area that developed or implemented the policy (chapter 8). Some

participants argued the merits of a separate body being established to conduct evaluations to ensure that evaluators were independent of policy and program agencies (for example, Tony Dreise et al., sub. 33). However, the Commission is of the view that agencies should be responsible for managing evaluations of the policies and programs that they design and implement to ensure that evaluations are integrated into the policy process and opportunities for learning are maximised. An appropriate level of independence can be ensured by requiring that evaluations be conducted by evaluators that are independent of the area of the agency with policy or program responsibility.

Participants recognised the trade-offs involved. For example, the Aboriginal Health Council of Western Australia said:

Independence and objectivity is essential for accurate information to be gathered through program evaluation. ... AHCWA [Aboriginal Health Council of Western Australia] recommends the framework should promote fair and balanced assessments in alignment with expected program outcomes. This could be achieved by engaging independent evaluators to complete reviews rather than commissioning bodies who may interpret outcomes in a way that supports a particular position or interest. (sub. 42, p. 3)

The Strategy does not endorse particular evaluation types, approaches or methods, but instead guides agencies and evaluators in adopting approaches and methods that are rigorous, fit-for-purpose and answer the questions that policy makers and Aboriginal and Torres Strait Islander people need answered. Many participants pointed out that the most appropriate methods and approaches will differ across evaluations, and that the Strategy should provide flexibility for agencies and evaluators to choose the best methods for each circumstance. The Victorian Aboriginal Child Care Agency, for example, said:

Approaches and methods that are appropriate for policies and programs affecting Aboriginal and Torres Strait Islanders should be based on a number of factors, including the purpose of the evaluation, ensuring there is a credible, robust design, and committing to development and delivery of a culturally endorsed and owned form and approach. (sub. 26, p. 4)

The National Indigenous Australians Agency also emphasised the importance of adopting fit-for-purpose methodologies:

It is important to allow space for tailoring evaluation methodologies developed through the strategy to the specifics of the policies and programs being evaluated. It is unlikely a one-size-fits-all approach will work to generate meaningful, relevant and useful findings for Indigenous communities as well as government. (sub. 86, p. 8)

Credibility is also about the processes used during evaluation planning and conduct to verify the results. Credibility of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people is enhanced by engaging with Aboriginal and Torres Strait Islander people at all stages of evaluation, clearly identifying limitations and having quality assurance measures in place (chapter 5).

Useful

This principle is about evaluations that are useful for decision makers and Aboriginal and Torres Strait Islander people alike. Evaluations that do not provide useful results are a waste of resources. When Australian Government agencies plan, commission or conduct an evaluation, the intention should always be to use the evaluation's findings to inform policy and program decisions. As Empowered Communities said:

Evaluation must be a living process, not one that produces unused evaluation products that sit on the shelves of policy makers. (sub. 41, p. 6)

Useful evaluations consider the needs of evaluation users in the planning phase (chapter 7). This includes policy makers as well as Aboriginal and Torres Strait Islander people in their many roles, including as decision makers, service providers and evaluation participants. When talking about current evaluation practice and its usefulness to Aboriginal and Torres Strait Islander service providers, the Victorian Aboriginal Child Care Agency said:

These evaluations overwhelmingly serve the needs of the government funding agency in that they are designed to respond to funding agency drivers of evaluation. While this does not mean that external evaluations are not useful for service delivery agencies, it does mean that often what is being evaluated does not necessarily reflect the evaluation priorities or what is important to Aboriginal people and organisations. (sub. 26, p. 5)

Participants also emphasised the need for evaluation to be useful to those participating in evaluations, reflecting the principle of reciprocity that is important in ethical frameworks for research with Aboriginal and Torres Strait Islander people. For example, the Smith Family said:

Evaluations must produce something that is useful for those participating in the evaluation (individuals, families, communities, and organisations) as well as for the commissioning Agency. (sub. 56, p. 9)

Usefulness, including from the viewpoint of Aboriginal and Torres Strait Islander people, is important when deciding which policies and programs to evaluate and what outcomes to measure. The Victorian Aboriginal Child Care Agency recommends:

That decisions about evaluation priorities consider the usefulness of the evaluation — in particular whether it will address research gaps and/or contribute to building an evidence base for Aboriginal programs. (sub. 26, p. 6)

Evaluations are most useful when they are of high quality and are timed to contribute to policy-making decisions. For high-quality evaluations, this means the right questions are asked, evaluators have engaged with those affected by the policy or program and evaluations are rigorous and impartial.

Evaluation should be embedded into the policy cycle so that planning and data collection can start early and evaluation questions be linked closely to policy and program objectives. It will also mean that evaluation findings are available at key decision points during

implementation, and support continuous quality improvement. The NSW Aboriginal Land Council, for example, advocated for one of the Strategy's principles to be:

Responsive: evaluation findings are used to inform program delivery and improve subsequent evaluations ... (sub. 91, p. 6)

Ethical

While it can be argued that ethical conduct is essential for evaluations to be both credible and useful, the Commission is proposing including 'ethical' as a separate principle to highlight its importance for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.

Many participants called — either explicitly or through the types of principles that they suggested be included in the Strategy — for the Strategy to be clear that evaluations should be conducted ethically. For example, the Lowitja Institute said:

The Strategy Principles should reflect a mix of guiding concepts to ensure ethical and high-quality evaluation in the unique context of program delivery and design for Aboriginal and Torres Strait Islander people. (sub. 50, p. 5)

And Aboriginal Affairs NSW said:

The framework should build from the guidelines and principles of ethical practice as a matter of course. ... The ethical clearance process ... requires that the impacts on communities and individuals are fully considered and methodologies are sound. (sub. 70, p. 2)

Indigenous Community Volunteers pointed out the benefits of ethical evaluation practices in improving quality:

The Strategy provides an opportunity to place greater emphasis on good practice evaluation. The rigour and credibility of an evaluation that cannot demonstrate good ethical practice is questionable. An evaluation with good ethical practice, including consideration and development of cultural capability, is far more likely to facilitate a safe environment where participants can provide an honest account. (sub. 88, p. 18)

Rather than define what ethical conduct looks like, the Commission considers that the Strategy should reference existing ethical guidelines for research involving Aboriginal and Torres Strait Islander people, including those developed by the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS 2012, 2019) and the National Health and Medical Research Council (NHMRC 2018a). A number of participants highlighted existing guidelines as being appropriate for use in evaluation. For example, the National Aboriginal Community Controlled Health Organisation said:

AIATSIS's (2012) and NHMRC's (2018) ethical guidelines for researchers working with Aboriginal and Torres Strait Islander people and communities are increasingly used to guide ethical conduct in evaluations, but should be adhered to more consistently by all government and non-government agencies. (sub. 95, p. 5)

There were also calls for the Strategy to clarify the circumstances in which formal ethical review was required in evaluations (for example, the Western Australian Government, sub. 74, p. 8).

Rather than be definitive, the Commission proposes that the Strategy require evaluations to include a systematic and well-documented assessment of ethical risks. The Australian Income Management Network highlighted the benefits of such an approach:

The value of being explicit and systematic in an approach to ethics is that this encourages open discussion about ethical concerns that may otherwise not occur. It can also prompt policy makers to consider ethical issues that fall outside of their personal values framework, or of the values framework built into the program logic. (sub. 15, p. 10)

While a more systematic approach to ethical conduct in evaluation is likely to increase the cost of evaluation, the benefits — that is, higher quality evaluations that Aboriginal and Torres Strait Islander people can trust are conducted in line with best practice ethical principles — are expected to outweigh the costs.

Recognising the cost and time implications of improving ethical conduct in evaluation, the Strategy should make clear that evaluations must ensure that sufficient resources are devoted to meeting ethical obligations, a point recognised by evaluation practitioners (for example, Indigenous Community Volunteers, sub. 88, p. 19).

Transparent

This principle recognises that transparency about evaluation conduct, processes and findings increases both accountability and learning (the key reasons for undertaking evaluations). As the Australian Institute of Aboriginal and Torres Strait Islander Studies said:

Publicly available evaluation reports are important for transparency and accountability of government spending, but also to build the knowledge and evidence base upon which Indigenous peoples and organisations can make decisions. (sub. 72, p. 15)

A lack of transparency — particularly failure to publish completed evaluation reports — was identified as a failing of current evaluation practices (chapters 3 and 7). There was widespread support from participants for all evaluation reports to be published, subject to privacy or cultural considerations (box 7.3).

As well as supporting learning and accountability, transparency allows evaluation users to judge the credibility and rigour of evaluation techniques used, and provides an incentive for agencies to commission and conduct high-quality evaluation:

... transparency and independence are crucial elements in ensuring evaluations are high quality and ultimately effective. (Tony Dreise et al., sub 33, p. 5)

Under the framework the process should be ethical and transparent, therefore the level of independence should not pose an issue. Any conflicts of interest should be declared and where possible minimised. (Aboriginal Health Council of South Australia, sub. 61, p. 4)

The principle of transparency also covers evaluation processes such as evaluation planning and commissioning, the importance of which was highlighted by some participants:

In light of the Commonwealth's Indigenous Procurement Policy we would like to see more accountability and transparency in how tendered contracts meet and support Indigenous expertise and capacity building. (Research School of Population Health (Australian National University), sub. 84, p. 3)

There needs to be transparent processes in regards to evaluation tenders and the criteria upon which evaluators are chosen. (National Family Violence Prevention Legal Services Forum, sub. 66, p. 4)

A maturity approach to implementing the principles

As noted in chapters 2 and 3, agencies are beginning from different points when it comes to building evaluation into the policy cycle, conducting and using high-quality evaluation and engaging meaningfully with Aboriginal and Torres Strait Islander people. Recognising progress, and learning from the successes of agencies that are further along the evaluation pathway, will be critical to support the implementation of the Strategy and promote long-term cultural change.

The Australian Evaluation Society advocated a 'maturity' approach to implementing the principles of the Strategy:

The Strategy should support different agencies at different levels of evaluation maturity to build an evaluative culture, and improve their capacity and capability to conduct, commission and manage evaluations of Indigenous policies and programs.

- The [Australian Public Service] and individual portfolios have different organisational cultures, histories and experience with evaluation, and in Indigenous affairs.
- The structures and processes adopted within different agencies should encourage and support staff to build evaluation maturity relevant to their organisation but aligned to the principle-based framework. (sub. 49, p. 14)

The benefits of a maturity approach for guiding government agencies in New Zealand to change their culture of engaging with Māori were also highlighted by Te Arawhiti (Office for Māori Crown Relations):

Culture change is difficult to achieve. ... Sometimes aspects of culture change can occur quickly, while others take years of investment and reinforcing. It will be important for agencies to adopt indicators of success, so they can know they are moving in the right direction. (2019c, p. 1)

A maturity approach is becoming more common in evaluation frameworks for Australian Government agencies. The IAS Evaluation Framework (DPMC 2018b) has a maturity matrix that sets out the practices required to meet a specified level of maturity against each of the Framework's eight principles (table 11.2).

Table 11.2 **Maturity matrix for selected principles from the IAS Evaluation Framework**

	<i>Beginning</i>	<i>Developing</i>	<i>Embedded</i>	<i>Leading</i>
Integrated	Awareness of the benefits of evaluation is low. Evaluation is seen as a compliance activity and threat. Fear of negative findings and recommendations leads to a perception of 'mandatory optimism' regarding program performance. Insufficient resources allocated to evaluation activities.	Appreciation of the benefits of evaluation improving. Evaluation is being viewed as core business for the Group, not simply a compliance activity.	A culture of evaluative thinking and continual improvement is embedded across the department, with lessons learnt being acted upon. Evaluation is seen as an integral component of sound performance management.	Evaluations motivate improvements in program design and policy implementation.
Respectful	Some attention to benefit of evaluation to community in some studies, particularly employment in conduct of evaluation and return of results.	Focus on benefit of evaluation to community increasing in frequency and range of benefits considered. Guidelines for collaboration with Indigenous people on design, conduct and use of evaluation findings are being integrated into usual practice.	Evaluation practice involves some collaboration with Indigenous people on aspects of design, conduct and use of evaluation findings.	Standard evaluation practice involves: collaboration with Indigenous people on all aspects of design, conduct and use of evaluation findings. Approaches are strengths-based, build partnerships and demonstrate cultural respect.
Transparent	Findings and recommendations held in program and policy areas. No follow up on the implementation of recommendations.	Findings and recommendations viewed as an opportunity to identify lessons learned. Evaluation findings and recommendations available in the research management system to improve the dissemination of lessons learned and inform policy development.	Findings widely disseminated and drive better performance. Website and guidance materials valuable resource for staff. Evaluation findings and reports are published where appropriate.	Findings are consistently used to optimise delivery and have influence outside the department. Materials provided to allow replication of evaluation results/findings. Evaluation reports are published in peer reviewed journals and presented at conferences.
Ethical	Ethics approval sought when necessary, depending on contextual requirements. Internal ethical checks on informed consent, and risk of harm.	Endorsement of ethical practice is sought as usual practice.	Ethical practice integrated into evaluation design, conduct and use.	Ethical practice meets the highest standards for respectful involvement of Indigenous Australians in evaluation design, conduct and use.

Source: ANAO (2019b, appendix 3).

The Department of Industry, Science, Energy and Resources also uses a maturity approach to improve and review its evaluation practices (DIIS 2019a; DIS 2015). When reviewing their strategy, the Department used the matrix to assess progress overall from ‘developing’ in 2015 to between ‘developing’ and ‘embedded’ by 2017 (DIS 2017, p. 5).

A maturity approach is also evident in several initiatives to improve research and engagement with Māori in New Zealand.

- Hudson et al. outlined a framework for Māori research ethics that identifies ‘progressive expectations of ethical behaviour’ for each of its principles, ranging from ‘minimum standards’ to ‘good practice’ to ‘best practice’ (2010, p. 4).
- Te Arawhiti (Office for Māori Crown Relations) used a maturity approach in its framework for improving the capability of New Zealand Government agencies in engaging with Māori. Levels of competency in the maturity matrix range from ‘unfamiliar’ to ‘comfortable’ to ‘confident’ to ‘capable’. The matrix is designed to be tailored to the needs and priorities of different agencies. It is ‘intended to guide agency leaders to assess where their agency is, determine where they want it to be and plan for how they will get there’ (2019c, p. 2).

The Commission sees merit in the Strategy using a maturity approach to give agencies a roadmap to improve evaluation practices. The practices and cultures for each level of maturity for the principles will be developed in more detail following the release of the draft Strategy with the aim of having a standard against which agencies’ progress in implementing the Strategy can be measured.

A Conduct of the project

The Commission has actively encouraged public participation in the development of the draft Indigenous Evaluation Strategy. This appendix outlines the engagement process undertaken and lists the organisations and individuals that have participated in this project.

The process was as follows:

- following receipt of the Letter of Direction on 10 April 2019, a circular was sent to identified interested parties
- an issues paper was released on 26 June 2019 to assist those wishing to make a submission to the project. Following the release of the issues paper, 112 submissions were received, of which three were conducted orally (table A.1). The submissions are available online at <https://www.pc.gov.au/inquiries/current/indigenous-evaluation/submissions#initial>. Three brief comments were also received
- consultations were held with a wide range of organisations and individuals, including Aboriginal and Torres Strait Islander community controlled organisations, land councils and peak bodies, Empowered Communities organisations and Leadership Group, Australian, state and territory government agencies, local government representatives, academics and evaluators. The Commission also consulted with government agencies, organisations and academics in New Zealand, and two Federal Government agencies in Canada (table A.2)
- the Commission held two roundtables. The first roundtable was held on 26 November 2019 on the topic of ‘An Indigenous Evaluation Strategy: Objectives, Principles and Defining Success’. The second roundtable was held on 28 November 2019 on the topic of ‘Culture, Capability and Governance in Government for Better Outcomes through Evaluation’ (table A.3)
- the Commission also presented at several forums, including at the Lowitja Institute International Indigenous Health and Wellbeing Conference in June 2019 and the National Aboriginal Community Controlled Health Organisation (NACCHO) National Members’ Conference in November 2019.

The Commission welcomes further submissions to discuss the content of the draft Strategy, background report and *A Guide to Evaluation under the Indigenous Evaluation Strategy*. Further information on making a submission can be found in the draft Strategy and on the Commission’s website. Submissions are due by 3 August 2020.

Table A.1	Submissions received ^a	
<i>Participants</i>		<i>Submission number</i>
Aboriginal & Torres Strait Islander Legal Service (QLD) Ltd		107
Aboriginal Affairs NSW		70
Aboriginal Carbon Foundation		111 #
Aboriginal Health Council of South Australia Limited (AHCSA)		61
Aboriginal Health Council of Western Australia (AHCWA)		42
Aboriginal Justice Caucus Working Group on Family Violence		106
Aboriginal Medical Services Alliance NT (AMSANT)		81
AbSec - NSW Child, Family and Community Peak Aboriginal Corporation		09
Accountable Income Management Network (AIMN)		15
AIME Mentoring		37
Allison, Fiona and LawRight		18
Altman, Jon		23
Anderson, Professor Clair		105 #*
Angelo, Denise; Hudson, Dr Catherine; Macqueen, Dr Susy		76
ARC Centre for Excellence for the Dynamics of Language		93
Audiology Australia		27
Australian Council of TESOL Associations (ACTA)		87
Australian Evaluation Society		49
Australian Government Department of Health		35 #
Australian Government Department of the Environment and Energy (DoEE)		101
Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)		72
Australian Institute of Health and Welfare (AIHW)		99
Bailey, James		29
Bielefeld, Dr Shelley		65
Bond, Dr Chelsea; Macoun, Dr Alissa; Staines, Dr Zoe; Brady, Ms Karla; Mukandi, Dr Bryan; Strakosch, Dr Elizabeth; Hassall, Ms Keryn; Singh, Dr David		40
Cape York Institute		69
Carers NSW		90
Cargo, Associate Professor Margaret		22 #
CatholicCare NT and The Australian Centre for Social Enterprise		75
Central Australian Aboriginal Congress		48 #
Central Australian Academic Health Science Network (CA AHSN)		85 #
Children's Ground		73 #
City of Ballarat		54
Clear Horizon		78
Close the Gap Campaign Steering Committee		71
Commissioner for Children and Young People (WA)		08
Congress of Aboriginal and Torres Strait Islander Nurses and Midwives (CATSiNaM)		01

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
Couzos, Associate Professor Sophia	92
Cox, Adjunct Professor Eva	68
Cross, Sherrie	82
Curran, Dr Liz	02
Darling Downs and West Moreton Primary Health Network	06
Dillon, Michael	16
Dreise, Professor Tony; Markham, Dr Francis; Dillon, Mr Michael; Hunt, Assoc Professor Janet; Dinku, Dr Yonatan; Weepers, Ms Jayne; Woods, Ms Kaely; Yap, Dr Mandy; Lahn, Dr Julie	33
Empowered Communities	41
Evans, Greg	110*
First Nations Media Australia	30
First Peoples Disability Network Australia	43
Generation One, Minderoo Foundation	05
GroundUP Northern Institute Charles Darwin University	62 #
Indigenous Allied Health Australia (IAHA)	31
Indigenous Business Australia (IBA)	36
Indigenous Community Volunteers (ICV)	88
Indigenous Data Network	32
Institute for Urban Indigenous Health (IUIH)	53
Jobs Australia	57
Just Reinvest NSW	80
Kimberley Aboriginal Law & Cultural Centre (KALACC)	103
Kimberley Community Legal Services (KCLS)	67
Literacy for Life Foundation	28
Local Government Association of Queensland (LGAQ)	102
Local Government Association of the Northern Territory (LGANT)	39
Lowitja Institute	50
Mangoyana, Robert	109
Minerals Council of Australia (MCA)	34
Moran, Mark	64
Name withheld – Oral submission	03
National Aboriginal and Torres Strait Islander Health Standing Committee (NATSIHSC)	104
National Aboriginal and Torres Strait Islander Leadership in Mental Health (NATSILMH) and The Centre of Best Practice in Aboriginal and Torres Strait Islander Suicide Prevention (CBPATSIISP)	63
National Aboriginal and Torres Strait Islander Legal Services (NATSILS)	97
National Aboriginal Community Controlled Health Organisation (NACCHO)	95
National Family Violence Prevention Legal Services Forum (FVPLS)	66
National Health and Medical Research Council (NHMRC)	89 #
National Health Leadership Forum (NHLF)	59

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
National Indigenous Australians Agency (NIAA)	86 #
National Justice Project	51
National Mental Health Commission	19
National Social Security Rights Network (NSSRN)	10
NSW Aboriginal Land Council	91 #
Nunkuwarrin Yunti of South Australia Inc	17
Office of Indigenous Education and Engagement (OIEE), RMIT University - Oral submission	98
PricewaterhouseCoopers (PwC) Indigenous Consulting Pty Ltd	58
Queensland Aboriginal and Islander Health Council (QAIHC)	55
Queensland Indigenous Family Violence Legal Service (QIFVLS)	25
Queensland Nurses and Midwives' Union (QNMU)	12
Reconciliation Australia	79
Regional Development Australia (Northern Territory Committee)	07 #
Regional Development Australia Fitzroy and Central West (RDAFCW)	13
Regional Development Australia Wheatbelt WA	24
Research School of Population Health (Australian National University)	84 #
Royal Australasian College of Physicians (RACP)	77
Save the Children	52
Smith, Professor James	47 #
Social Ventures Australia (SVA)	83
Sullivan, Dr Patrick; Hunt, Dr Janet; Lahn, Dr Julie	11 #
Tasmanian Government	100
Tauondi Aboriginal Corporation (trading as Tauondi Aboriginal College)	94
The Fred Hollows Foundation (TFHF)	14
The Smith Family	56
The University of Queensland	20
Torres Shire Council	108
Tranby National Indigenous Adult Education and Training	96
Victorian Aboriginal Child Care Agency (VACCA)	26
Victorian Aboriginal Community Controlled Health Organisation (VACCHO)	44
Walgett Aboriginal Medical Service (WAMS)	04
Walter, Professor Maggie - Oral submission	112
Western Australian Council of Social Services (WACOSS)	21
Western Australian Government	74 #
Williams, Dr Megan; Ragg, Dr Mark	45
Winnunga Nimmityjah AHCS	46
Yothu Yindi Foundation (YYF)	60 #
Yuwaya Ngarra-li partnership, Dharriwaa Elders Group and UNSW	38 #*

^a An asterisk (*) indicates that the submission contains confidential material NOT available to the public. A hash (#) indicates that the submission includes attachments.

Table A.2 Consultations

Participants

ACT

Aboriginal Hostels Ltd
ACT Chief Minister, Treasury and Economic Development Directorate
Attorney-General's Department
Australian Bureau of Statistics
Australian Indigenous Governance Institute
Australian Institute of Aboriginal and Torres Strait Islander Studies
Australian Institute of Health and Welfare
Australian National Audit Office
Australian Public Service Commission
Australian Government Treasury (The Treasury)
Centre for Social Research and Methods, Australian National University
Department of Agriculture
Department of Communications and the Arts
Department of Education
Department of Employment, Skills, Small & Family Business
Department of Finance
Department of Foreign Affairs and Trade
Department of Health
Department of Industry, Innovation and Science
Department of Infrastructure, Transport, Cities and Regional Development
Department of Social Services
Department of the Environment and Energy
Department of the Prime Minister and Cabinet
Dreise, Professor Tony
Indigenous Allied Health Australia
Indigenous Business Australia
National Aboriginal Community Controlled Health Organisation
National Indigenous Australians Agency
Office of the Registrar of Indigenous Corporations
Plowman, Colin
Winnunga Nimmityjah Aboriginal Health and Community Services
Yap, Dr Mandy
Yu, Eunice

(continued next page)

Table A.2 (continued)

Participants

New South Wales

Aboriginal Affairs New South Wales
Aboriginal Health and Medical Research Council
AbSec New South Wales Child, Family and Community Peak Aboriginal Corporation
Australian Evaluation Society
Australian Human Rights Commission
Barang Regional Alliance
Empowered Communities Central Team
Empowered Communities Leadership Group
Fetterman, Professor David
First Peoples Disability Network
Hudson, Sara
Inner Sydney Empowered Communities
Lester, Danny, New South Wales Deputy Ombudsman
New South Wales Aboriginal Land Council
New South Wales Department of Education
New South Wales Department of Premier and Cabinet
New South Wales Ministry of Health
New South Wales Treasury
Oscar, Commissioner June, Aboriginal and Torres Strait Islander Social Justice Commissioner
Wehipeihana, Nan

Northern Territory

Aboriginal Peak Organisations Northern Territory
Menzies School of Health Research
Northern Territory Department of the Chief Minister
Northern Territory Department of Treasury and Finance

South Australia

Aboriginal Health Council of South Australia
Aboriginal Legal Rights Movement
Fry, Eddie, Chair, Indigenous Business Australia; Chair, Indigenous Land and Sea Corporation
South Australian Aboriginal Advisory Council
South Australian Department of the Premier and Cabinet
South Australian Department of Treasury and Finance
South Australian Health and Medical Research Institute

Tasmania

Department of Communities Tasmania
Tasmanian Aboriginal Centre
Tasmanian Department of Premier and Cabinet
Tasmanian Department of Treasury and Finance
Walter, Professor Maggie

(continued next page)

Table A.2 (continued)

Participants

Queensland

Andrews, Ross, Mayor, Yarrabah Aboriginal Shire Council
Cape York Partnerships
Department of Foreign Affairs and Trade, Torres Strait Treaty Liaison Office
Gur A Baradharaw Kod Torres Strait Sea and Land Council
Institute for Urban Indigenous Health
Local Government Association of Queensland
Mills, Donella, Chair, National Aboriginal Community Controlled Health Organisation
Moran, Professor Mark
Malone, Vonda, Mayor, Torres Shire Council
National Aboriginal and Torres Strait Islander Higher Education Consortium
Queensland Aboriginal and Islander Health Council
Queensland Aboriginal and Torres Strait Islander Child Protection Peak
Queensland Department of Aboriginal and Torres Strait Islander Partnerships
Queensland Department of the Premier and Cabinet
Queensland Government Statistician's Office
Queensland Indigenous Family Violence Legal Service
Queensland Productivity Commission
Queensland Treasury
Stronger Smarter Institute
Torres Strait Regional Authority

Victoria

Academy of Sport, Health and Education, Shepparton
Commission for Children and Young People Victoria
Gruen, Nicholas, Lateral Economics
Kaiela Arts
Kaiela Institute
Langton, Professor Marcia
National Aboriginal and Torres Strait Islander Legal Services
Rogers, Professor Patricia and Peersman, Associate Professor Greet, BetterEvaluation
Rose, Dr James
Rumbalara Football Netball Club
Secretariat of National Aboriginal and Islander Child Care
Victorian Aboriginal Child Care Agency
Victorian Aboriginal Children's Forum
Victorian Department of Health and Human Services
Victorian Department of Premier and Cabinet
Victorian Department of Treasury and Finance

(continued next page)

Table A.2 (continued)

Participants

Western Australia

Aarnja

Aboriginal Health Council of Western Australia

Aboriginal Legal Services of Western Australia

Binarri-binyja Yarrawoo

Broome Regional Aboriginal Medical Services

Drew, Professor Neil, Australian Indigenous HealthInfoNet

Dudgeon, Professor Pat

Kimberley Aboriginal Law and Cultural Centre

Kimberley Aboriginal Medical Services

Kimberley Land Council

KRED Enterprises

Majarlin Kimberley Centre for Remote Health, University of Notre Dame

Miriuwung Gajerrong (MG) Corporation

Nulungu Research Institute, University of Notre Dame

Nyamba Buru Yawuru

Walker, Adjunct Associate Professor Roz

Western Australian Department of the Premier and Cabinet

Western Australian Treasury

Westerman, Adjunct Professor Tracy

Wunan Foundation

Canada

Privy Council Office

Treasury Board Secretariat

New Zealand

Cram, Dr Fiona

Kukutai, Professor Tahu

Mā te Rae (the Māori Evaluation Association)

Masters-Awatere, Dr Bridgette

Moewaka Barnes, Professor Helen, SHORE & Whariki Research Centre

Te Puni Kōkiri (the Ministry of Māori Development)

The Independent Māori Statutory Board, Auckland

The Ministry of Social Development

The New Zealand Treasury

The Social Investment Agency

Wolfgramm, Tania

Table A.3 Roundtable participants

26 November 2019 – ‘An Indigenous Evaluation Strategy: Objectives, principles and defining success’ – Canberra

Australian Human Rights Commission
Australian Indigenous Doctors’ Association
Australian Institute of Family Studies
Australian Institute of Health and Welfare
Bond, Associate Professor Chelsea
Department of Finance
Department of the Prime Minister and Cabinet
Dreise, Professor Tony
Empowered Communities Central Team
Empowered Communities Leadership Group
Indigenous Allied Health Australia
Lowitja Institute
National Aboriginal and Torres Strait Islander Health Worker Association
National Indigenous Australians Agency
Walter, Professor Maggie
Yap, Dr Mandy

28 November 2019 – ‘Culture, capability and governance in government for better outcomes through evaluation’ – Canberra

Aboriginal Affairs New South Wales
Australian Government Treasury (The Treasury)
Australian Human Rights Commission
Australian Institute of Aboriginal and Torres Strait Islander Studies
Australian Institute of Health and Welfare
Australian National Audit Office
Coalition of Aboriginal and Torres Strait Islander Peak Organisations – New South Wales Aboriginal Land Council
Coalition of Aboriginal and Torres Strait Islander Peak Organisations – Northern Land Council
Department of Finance
Department of Health
Department of Industry, Innovation and Science
Department of Social Services
Department of the Prime Minister and Cabinet
National Indigenous Australians Agency
Peersman, Associate Professor Greet

References

- Aboriginal Affairs NSW 2018, *OCHRE: Five Years On*, Aboriginal Affairs NSW, Sydney.
- Aboriginal Affairs (NSW) and Department of Education (NSW) 2015, *Aboriginal Identification in NSW: the Way Forward*, Department of Education (NSW), Sydney.
- ABS (Australian Bureau of Statistics) 2010, *4703.0 - Framework for Measuring Wellbeing: Aboriginal and Torres Strait Islander Peoples*, <https://www.abs.gov.au/ausstats/abs@.nsf/mf/4703.0> (accessed 6 January 2020).
- 2013, *Perspectives on Aboriginal and Torres Strait Islander Identification in Selected Data Collection Contexts*, Information paper, Cat. No. 4726.0, Canberra.
- 2014, *Indigenous Status Standard*, Cat. No. 1200.0.55.008, Canberra.
- 2018, *Census of Population and Housing: Understanding the Increase in Aboriginal and Torres Strait Islander Counts, 2016*, Cat. No. 2077.0, Canberra.
- 2019, *Multi-Agency Data Integration Project (MADIP) Research Projects*, <https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+MADIP+Research+Projects> (accessed 13 January 2020).
- ACF and ANA (Administration for Children and Families and Administration for Native Americans) 2019, *FY 2016 Report to Congress on Outcome Evaluations of Administration for Native Americans Projects*.
- ACIL Allen Consulting 2014, *Evaluation of the Aboriginal and Torres Strait Islander Education Action Plan 2010-2014*, Final Evaluation Report, Report to the Aboriginal and Torres Strait Islander Education Advisory Group of the Education Council, Brisbane.
- 2018, *Evaluation of the Home Interaction Program for Parents and Youngsters: Final Report*, Canberra.
- ACT Chief Minister's Department 2010, *ACT Government Evaluation Policy and Guidelines*, December.
- AEA (American Evaluation Association) 2018, *Guiding Principles for Evaluators*, Washington.
- AES (Australasian Evaluation Society) 2013a, *Evaluators' Professional Learning Competency Framework*, Melbourne.
- 2013b, *Guidelines for the Ethical Conduct of Evaluations*, Melbourne.
- 2018, *Submission to the Independent Review of the Australian Public Service*, Melbourne.

-
- AGS (Australian Government Solicitor) 2013, *Public Governance, Performance and Accountability Act 2013 - What are the Key Changes for Commonwealth Authorities*, Fact Sheet 33, Canberra.
- AHA (Australian Healthcare Associates) 2018, *Evaluation of the Partners in Culturally Appropriate Care Program: Final Evaluation Report*, prepared for the Department of Health, Department of Health, Melbourne.
- AHRC (Australian Human Rights Commission) 2010, *Community Guide to the UN Declaration on the Rights of Indigenous Peoples*.
- 2012, *Access and Equity: Inquiry into the Responsiveness of Australian Government Services to Australia's Culturally and Linguistically Diverse Population*, Australian Human Rights Commission Submission to the Access and Equity Inquiry Panel, Sydney.
- AIATSIS (Australian Institute of Aboriginal and Torres Strait Islander Studies) 2012, *Guidelines for Ethical Research in Australian Indigenous Studies*, Canberra.
- 2019, *2019 revision of the AIATSIS Guidelines for ethical research in Australian Indigenous Studies: Consultation draft*, Canberra.
- 2020, *Indigenous Research Exchange*, <https://aiatsis.gov.au/exchange> (accessed 30 April 2020).
- AIFS (Australian Institute of Family Studies) 2013, *Demystifying ethical review: CFCA resource sheet*, <https://aifs.gov.au/cfca/publications/demystifying-ethical-review> (accessed 4 February 2020).
- 2014, *Developing a culture of evaluation and research*, Child Family Community Australia (CFCA) Paper, 28, Australian Institute of Family Studies, Melbourne.
- 2015, *Empowerment evaluation: CFCA practitioner resource*, <https://aifs.gov.au/cfca/publications/empowerment-evaluation> (accessed 30 April 2020).
- 2017, *How to develop a program evaluation plan*, <https://aifs.gov.au/cfca/expert-panel-project/program-planning-evaluation-guide/evaluate-your-program-or-service/how-develop-evaluation-plan> (accessed 17 February 2020).
- nd, *Closing the Gap Clearinghouse*, <https://aifs.gov.au/projects/closing-gap-clearinghouse> (accessed 13 September 2019).
- AIHW and AIFS (Closing the Gap Clearinghouse, Australian Institute of Health and Welfare and Australian Institute of Family Studies) 2013, *What Works to Overcome Indigenous Disadvantage: Key Learnings and Gaps in the Evidence 2011-12*, Canberra.
- AIHW (Australian Institute of Health and Welfare) 2010, *National Best Practice Guidelines for Collecting Indigenous Status in Health Data Sets*, IHW 29, Canberra.
- 2017a, *Closing the Gap — A clearinghouse for research and evaluation evidence on what works to overcome Indigenous disadvantage*, <https://webarchive.nla.gov.au/awa/20170816053059/http://www.aihw.gov.au/closingthegap/> (accessed 13 September 2019).

-
- 2017b, *Trends in Indigenous Mortality and Life Expectancy 2001-2015: Evidence from the Enhanced Mortality Database*, Canberra.
- 2020, *Review of the Two National Indigenous Specific Primary Health Care Datasets: The Online Services Report and the National Key Performance Indicators*, Canberra.
- Alexander, E. and Thodey, D. 2018, *Independent Review into the Operation of the Public Governance, Performance and Accountability Act 2013 and Rule*, Canberra.
- Alliance for Useful Evidence 2016, *Using Research Evidence: A Practice Guide*, Nesta Foundation, London.
- Al-Yaman, F. and Higgins, D. 2011, *What Works to Overcome Indigenous Disadvantage: Key Learnings and Gaps in the Evidence 2009-10*, Closing the Gap Clearinghouse, Canberra.
- Ames, P. and Wilson, J. 2016, *Unleashing the Potential of Randomised Controlled Trials in Australian Governments*, 55, M-RCBG Associate Working Paper Series, Mossavar-Rahmani Center for Business and Government, Harvard Kennedy School, Cambridge.
- ANA (Administration for Native Americans) 2019a, *Evaluation*, <https://www.acf.hhs.gov/ana/research> (accessed 24 April 2020).
- 2019b, *Team*, <https://www.acf.hhs.gov/ana/team> (accessed 11 May 2020).
- 2019c, *What is an ANA impact visit?*, <https://www.acf.hhs.gov/ana/resource/what-is-an-ana-impact-visit-0> (accessed 15 March 2019).
- nd, *About SEDS*, <https://www.acf.hhs.gov/ana/programs/seds/about> (accessed 24 April 2020).
- ANAO (Australian National Audit Office) 2003, *Annual Performance Reporting*, Performance Audit, No. 11 2003-04, Canberra.
- 2007, *Application of the Outcomes and Outputs Framework*, Performance Audit, No. 23 2006-07, Canberra.
- 2011a, *AuditFocus*, Canberra.
- 2011b, *Development and Implementation of Key Performance Indicators to Support the Outcomes and Programs framework*, Performance Audit, No. 5 2011-12, Canberra.
- 2012, *Australian Government Coordination Arrangements for Indigenous Programs*, Performance Audit, No. 8 2012-13, Canberra.
- 2013a, *Indigenous Early Childhood Development: Children and Family Centres*, Performance Audit, No. 38 2012-13, Canberra.
- 2013b, *The Australian Government Performance Measurement and Reporting Framework: Pilot Project to Audit Key Performance Indicators*, Performance Audit, No. 28 2012-13, Canberra.
- 2014, *Initiatives to Support the Delivery of Services to Indigenous Australians*, Performance Audit, No. 45 2013-14, Canberra.
-

-
- 2015, *Administration of the Indigenous Legal Assistance Program*, Performance Audit, No. 22 2014-15, Canberra.
 - 2017a, *A guide to conducting performance audits*, <https://www.anao.gov.au/work/corporate/guide-to-conducting-performance-audits> (accessed 23 March 2020).
 - 2017b, *Indigenous Advancement Strategy*, Performance Audit, No. 35 2016-17, Canberra.
 - 2017c, *Indigenous Aged Care*, Performance Audit, No. 53 2016-17, Canberra.
 - 2018, *The Implementation and Performance of the Cashless Debit Card Trial*, Canberra.
 - 2019a, *Closing The Gap*, Performance Audit, No. 27 2018-19, Canberra.
 - 2019b, *Evaluating Aboriginal and Torres Strait Islander Programs*, Performance Audit, 47 2018-19, Canberra.
 - 2020, *Aboriginal and Torres Strait Islander Participation Targets in Major Procurements*, Performance Audit, No. 25 2019-20, Canberra.
 - Anaya, J. 2012, *Report of the Special Rapporteur on the Rights of Indigenous Peoples, James Anaya — Addendum: The Situation of Indigenous Peoples in the United States of America*, A/HRC/21/47/Add.1, United Nations General Assembly.
 - 2014, *Report of the Special Rapporteur on the Rights of Indigenous Peoples, James Anaya — Addendum: The Situation of Indigenous Peoples in Canada*, A/HRC/27/52/Add.2, United Nations General Assembly.
 - Anderson, V. 2017, 'Criteria for evaluating qualitative research', *Human Resource Development Quarterly*, vol. 28, no. 2, pp. 125–133.
 - ANZEA (Aotearoa New Zealand Evaluation Association) 2011, *Evaluator Competencies*.
 - and Superu (Aotearoa New Zealand Evaluation Association and Social Policy Evaluation and Research Unit) 2015, *Evaluation standards for Aotearoa New Zealand*.
 - ANZSOG (Australia and New Zealand School of Government) 2018, *Why Evaluation Matters: Five Mistakes to Avoid When Starting a New Program*, <https://www.anzsog.edu.au/resource-library/news-media/why-evaluation-matters> (accessed 12 May 2020).
 - APSC (Australian Public Service Commission) 2015, *Aboriginal and Torres Strait Islander Cultural Capability: A Framework for Commonwealth Agencies*, Performance Audit, Canberra.
 - 2019, *APS HR Professional Stream Strategy A Structured Way to Build and Uplift Strategic HR Skills across the APS*, Canberra.
 - 2020a, *2021 Graduate Program*, <https://www.apsc.gov.au/hr-professional-stream/2021-graduate-program> (accessed 15 April 2020).
 - 2020b, *A structured way to build and uplift strategic HR skills across the APS*, <https://www.apsc.gov.au/aps-hr-professional-stream> (accessed 5 July 2020).

-
- Atkinson, J. 2002, 'Dadirri: Listening to one another', *Trauma Trails: Recreating Song Lines*, ProQuest Ebook Central, pp. 6–22.
- Auditor-General 1991a, *Evaluation in the Preparation of the Budget*, Efficiency Audit, No. 13 1991-92, Canberra.
- 1991b, *Implementation of Program Evaluation — Stage 1*, No. 23 1990-91, Canberra.
- 1997, *Program Evaluation in the Australian Public Service*, Canberra.
- Australian Government 2020a, *AusTender Advanced Search: Contract Notices*, <https://www.tenders.gov.au/Search/Home> (accessed 14 January 2020).
- 2020b, *Closing the Gap Report 2020*, Canberra.
- Australian Health Ministers' Advisory Council 2006, *National Aboriginal and Torres Strait Islander Health Data Principles*, COAG Health Council.
- Baily, R., Hardie-Boys, N., Bailie, J., Carr, J., Moodie, D., McLean, G., Scanlen, A., Bainbridge, R., Monaghan, R., Walke, E. and Bailie, R. 2018, *Evaluation of the Australian Government's Investment in Aboriginal and Torres Strait Islander Primary Health Care through the Indigenous Australians' Health Program: Monitoring and Evaluation Design Report*, Allen + Clarke, Wellington.
- Bamberger, M. 2012a, *Annex 1: Strengths and Weaknesses of Quantitative Evaluation Designs*, <https://archive.interaction.org/resources/training/annex-1-strengths-and-weaknesses-quant-evaluation-approaches> (accessed 17 February 2020).
- 2012b, *Annex 2: Strengths and Weaknesses of Qualitative Evaluation Designs*, <https://archive.interaction.org/resources/training/annex-2-strengths-and-weaknesses-qualitative-evaluation-designs> (accessed 17 February 2020).
- Banks, G. 2009, 'Evidence-based policy-making: What is it? How do we get it?', *Challenges of Evidence-based Policy-Making*, Productivity Commission, Canberra.
- 2013, 'Introduction - why this roundtable?', *Better Indigenous Policies: The Role of Evaluation*, Productivity Commission, Canberra, pp. 17–20.
- Beal, S.J. and Kupzyk, K.A. 2014, 'An introduction to propensity scores: What, when, and how', *Journal of Early Adolescence*, vol. 34, no. 1, pp. 66–92.
- Better Evaluation 2011, *Utilization-Focused Evaluation*, https://www.betterevaluation.org/en/plan/approach/utilization_focused_evaluation (accessed 19 March 2020).
- 2014, *Cost-effectiveness analysis*, <https://www.betterevaluation.org/evaluation-options/CostEffectivenessAnalysis> (accessed 10 January 2019).
- 2016, *Cultural Competency*, https://www.betterevaluation.org/en/evaluation-options/cultural_competence_evaluation (accessed 2 July 2020).
- BHI (Bureau of Health Information) 2015, *Technical supplement: Adult Admitted Patient Survey, 2014*, Sydney.
- 2016, *Patient perspectives: Hospital care for Aboriginal people*, Sydney.

-
- Biddle and AIFS (Biddle and Australian Institute of Family Studies) 2014, *Data about and for Aboriginal and Torres Strait Islander Australians*, Closing the Gap Clearinghouse, Canberra.
- Blomkamp, E. 2018, 'The promise of co-design for public policy', *Australian Journal of Public Administration*, vol. 77, no. 4, pp. 729–743.
- Blöndal, J.R., Bergvall, D., Hawkesworth, I. and Deighton-Smith, R. 2008, 'Budgeting in Australia', *OECD Journal on Budgeting*, vol. 8, no. 2, pp. 1–64.
- Bond, C., Holder, R., Jeffries, S. and Fleming, C. 2017, *Evaluation of the Specialist Domestic and Family Violence Court Trial in Southport*, Griffith Criminology Institute, Griffith University.
- Bourassa, C. 2017, 'Research as Noojimo Mikana: Creating Pathways to Culturally Safe Care for Indigenous People', presented at 6th Annual NHMRC Symposium on Research Translation, Brisbane, 15 November, <https://webcast.gigtv.com.au/Mediasite/Play/5bbb6b4e399d4e5782cffe5be546f1891d?catalog=a6293e7e704d44f7a7f41fa63511db6e21> (accessed 29 April 2020).
- Bowes, J. and Grace, R. 2014, *Review of Early Childhood Parenting, Education and Health Intervention Programs for Indigenous Children and Families in Australia*, Issues Paper no. 8, Closing the Gap Clearinghouse.
- Brackertz, N., Davidson, J. and Wilkinson, A. 2017, *How Can Aboriginal housing in NSW and the Aboriginal Housing Office provide the best opportunity for Aboriginal People?*, Australian Housing and Urban Research Institute.
- Bradwell, P. and Marr, S. 2008, *Making the most of collaboration: an international survey of public service co-design*, Report 23, Price Waterhouse Coopers, London.
- Bray, J.R., Gray, M. and 't Hart, P. 2019, *Evaluation and Learning from Failure and Success*, Australia and New Zealand School of Government.
- Breckon, J. and Mulgan, G. 2015, *Better Public Services Through Experimental Government*, Alliance for Useful Evidence, United Kingdom.
- Brouwer, R. and Georgiou, S. 2012, 'Economic evaluation', *Animal Waste, Water Quality and Human Health*, World Health Organisation, London.
- Bruhn, J. 2014, 'Identifying useful approaches to the governance of indigenous data', *The International Indigenous Policy Journal*, vol. 5, no. 2.
- Cabaj, M. 2017, *Weaving Together Accountability and Learning in Complex Situations*, 22 December.
- Cabinet Office (UK) 2018, *What Works Network - Membership Criteria*.
- Cairney, S., Abbott, T., Quinn, S., Yamaguchi, J., Wilson, B. and Wakerman, J. 2017, 'Interplay wellbeing framework: a collaborative methodology "bringing together stories and numbers" to quantify Aboriginal cultural values in remote Australia', *International Journal for Equity in Health*, vol. 16, no. 68, pp. 1–13.

-
- Campbell Collaboration 2020, *History of the Campbell Collaboration*, <https://campbellcollaboration.org/about-campbell/history.html> (accessed 30 March 2019).
- Campbell, S. and Harper, G. 2012, *Quality in Policy Impact Evaluation: Understanding the Effects of Policy from Other Influences (Supplementary Magenta Book Guidance)*, December, United Kingdom.
- Carlson, T., Moewaka Barnes, H. and McCreanor, T. 2017, *Kaupapa Māori Evaluation: A Collaborative Journey*, New Zealand Council for Educational Research.
- CDC (Centers for Disease Control and Prevention) nd, *Overview of Tribal Epidemiology Centers*.
- Centre for Epidemiology and Evidence 2017, *Developing and using a program logic: A guide*, Evidence and Evaluation Guidance Series, Population and Public Health Division, NSW Ministry of Health, Sydney.
- 2019, *Commissioning Economic Evaluations: A Guide*, Evidence and Evaluation Guidance Series, Population and Public Health Division, NSW Ministry of Health, Sydney.
- Centre of Evaluation Excellence 2015, *2013 Annual Report on the Health of the Evaluation Function*, <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/2013-annual-report-health-evaluation-function.html> (accessed 4 May 2020).
- CES (Canadian Evaluation Society) 2018, *Competencies for Canadian Evaluation Practice*, Renfrew, Ontario.
- Charles Darwin University and NT Government 2017, *Indigenous Evaluation Research*, <http://recier.cdu.edu.au/indigenous-evaluation-research/> (accessed 10 August 2019).
- Chouinard, J.A. and Cousins, J.B. 2007, 'Culturally competent evaluation for Aboriginal communities: A review of the empirical literature', *Journal of MultiDisciplinary Evaluation*, vol. 4, no. 8, pp. 40–57.
- CLC (Central Land Council) nd, *Indigenous Ecological Knowledge | Central Land Council, Australia*, <https://www.clc.org.au/index.php?/articles/cat/indigenous-ecological-knowledge/> (accessed 21 May 2020).
- Closing the Gap Clearinghouse 2013, *Strategies and Practices for Promoting the Social and Emotional Wellbeing of Aboriginal and Torres Strait Islander People*, Resource sheet 19.
- COAG (Council of Australian Governments) 2009, *National Indigenous Reform Agreement (Closing the Gap)*.
- 2018, *COAG Statement on the Closing the Gap Refresh 12 December 2018*.
- Coalition of Aboriginal and Torres Strait Islander Peak Organisations 2018, *A new way of working: Talking about what's needed to close the gap in life outcomes between Aboriginal and Torres Strait Islander people and other Australians*, Canberra.

-
- 2020, *Community engagement snapshots: A summary of what Aboriginal and Torres Strait Islander people and communities said about the new National Agreement on Closing the Gap*, Canberra.
- Cobb-Clark, D. 2013, 'The case for making public policy evaluations public', *Better Indigenous Policies: The Role of Evaluation*, Productivity Commission, Canberra, pp. 81–92.
- Cochrane 2020, *About us*, <https://www.cochrane.org/about-us> (accessed 30 March 2019).
- Coglianesse, C. 2012, *Evaluating the Impact of Regulation and Regulatory Policy*, OECD, Paris.
- Conway, J., Lawn, S., Crail, S. and McDonald, S. 2018, 'Indigenous patient experiences of returning to country: a qualitative evaluation on a Country Health SA Dialysis Bus', *BMC Health Services Research*, vol. 18, no. 1010, pp. 1–18.
- Cortis, N., Smyth, C., Wade, C. and Katz, I. 2019, 'Changing practice cultures in statutory child protection: Practitioners' perspectives', *Child and Family Social Work*, vol. 24, no. 1, pp. 50–58.
- Cram, F. 2009, 'Maintaining indigenous voices', *The Handbook of Social Research Ethics*, pp. 308–322.
- , Kataraina, P. and Paipa, K. 2018, 'Kaupapa Māori Evaluation in Aotearoa New Zealand', *New Directions for Evaluation*, vol. 159, pp. 63–77.
- Crowley, M., Dodge, K.A., Barnett, W.S., Phaedra, C., Duffy, S., Graham, P., Greenberg, M., Haskins, R., Hill, L., Jones, D.E., Karoly, L.A., Kuklinksi, M.R. and Plotnick, R. 2018, 'Standards of evidence for conducting and reporting economic evaluations', *Prevention Science*, no. 19, pp. 366–390.
- Crown-Indigenous Relations and Northern Affairs Canada 2019a, *Pre-Transformation Reports of the Audit and Evaluation Sector*, <https://www.rcaanc-cirnac.gc.ca/eng/1536848932394/1536848987127> (accessed 28 April 2020).
- 2019b, *Self-government*, <https://www.rcaanc-cirnac.gc.ca/eng/1100100032275/1529354547314#chp1> (accessed 30 April 2020).
- CSRE (Centre for Social Research and Evaluation (Te Pokapu Rangahaū Arotake Hapori)) 2004, *Nga Ara Tohutohu Rangahau Māori - Guidelines for Research and Evaluation with Māori*, Ministry of Social Development, Wellington.
- Davis, M. and Holcombe, S. 2010, "'Whose Ethics?': Codifying and Enacting Ethics in Research Settings', *Australian Aboriginal Studies*, no. 2, pp. 1–9.
- DAWR (Department of Agriculture and Water Resources) 2017, *Evaluation Policy*, Canberra.
- Deaton, A. and Cartwright, N. 2018, 'Understanding and misunderstanding randomised controlled trials', *Social Science and Medicine*, vol. 210, pp. 2–21.
- Deloitte Access Economics 2017, *Cost Benefit Analysis of the Murri School Healing Program*, Healing Foundation, Canberra.

-
- Denzin, N. 1989, *Interpretive Interactionism*, Sage, California.
- Department of Aboriginal Affairs 1981, *Annual Report 1980-81*, Parliamentary Paper 302/1981, Canberra.
- 1983, *Annual Report: 1982-83*, Parliamentary Paper 260/1983, Canberra.
- Department of Education (NSW) 2019, *Data types - strengths and limitations*, <https://education.nsw.gov.au/teaching-and-learning/professional-learning/evaluation-resource-hub/turning-data-into-evidence/data-types-strengths-and-limitations> (accessed 5 April 2020).
- 2020, *Evaluative thinking*, <https://education.nsw.gov.au/teaching-and-learning/professional-learning/evaluation-resource-hub/evaluative-thinking> (accessed 13 March 2020).
- Department of Finance 1994a, *Doing Evaluations: A practical guide*, Canberra.
- 1994b, *The Use of Evaluation in the 1994-95 Budget*, Canberra.
- 2017, *Annual Performance Statements for Commonwealth Entities*, Resource Management Guide No. 134.
- 2019, *Commonwealth Procurement Rules*, Canberra.
- Department of Justice Canada 2018, *Principles: Respecting the Government of Canada's Relationship with Indigenous Peoples*.
- Department of the Environment 2015, *Evaluation Policy 2015-2020*, Canberra.
- Department of Treasury and Finance (VIC) 2013, *Economic Evaluation for Business Cases Technical Guidelines*, Victorian Government, Melbourne.
- Department of Treasury (WA) 2015, *Evaluation Guide*, Program Evaluation Unit, Perth, <http://www.programevaluation.wa.gov.au/ProgramEvaluationWA/media/Documents/Tools-and-templates/Program-Evaluation-Guide.pdf> (accessed 20 September 2019).
- DESSFB (Department of Employment, Skills, Small and Family Business) 2020, *The Evaluation of Job Services Australia 2012 - 2015*, Canberra.
- DFAT (Department of Foreign Affairs and Trade) 2017a, *DFAT Aid Evaluation Policy*, Canberra.
- 2017b, *DFAT Monitoring and Evaluation Standards*, Canberra.
- 2017c, *Evaluation in DFAT: Improving Quality and Use*, Canberra.
- DHHS (Department of Health and Human Services (VIC)) 2016, *Improving Cultural Responsiveness of Victorian Hospitals*, Melbourne.
- 2017, *Cradle to Kinder*, Melbourne.
- Di Francesco, M. 1996, *Program Evaluation and Policy Management in Australian Central Agencies*, December, Australian National University.
- DIIS (Department of Industry, Innovation and Science) 2019a, *Evaluation Strategy 2017 to 2021*, Canberra.

-
- 2019b, *Tradex Monitoring Evaluation Report 2018*.
- Dillon, M.C. 1992, 'Program Evaluation and Monitoring in Aboriginal and Torres Strait Islander Affairs: A Strategic Overview', *Program evaluation and performance monitoring: an Australian perspective*, Macmillan, South Melbourne, p. 189.
- DIRD (Department of Infrastructure and Regional Development) 2016, *Evaluation Resource Booklet 2016–2021*, Canberra.
- DIS (Department of Industry and Science) 2015, *Evaluation Strategy 2015-2019*, Canberra.
- 2017, *Evaluation Strategy: Post Commencement Review*, Canberra.
- DISER (Department of Industry, Science, Energy and Resources) 2017, *Business Longitudinal Analysis Data Environment (BLADE)*, <https://www.industry.gov.au/data-and-publications/business-longitudinal-analysis-data-environment-blade> (accessed 30 April 2020).
- DJSB (Department of Jobs and Small Business) 2018, *Relocation Assistance to Take Up a Job Evaluation Report*, Canberra.
- 2019, *Job Commitment Bonus for Young Australians Evaluation*, Canberra.
- Dockett, S., Perry, B. and Kearney, E. 2010, *School Readiness: What Does it Mean for Indigenous Children, Families, Schools and Communities*, Issues Paper 2, Closing the Gap Clearinghouse.
- DOFD (Department of Finance and Deregulation) 2010, *Strategic Review of Indigenous Expenditure*, Canberra.
- 2012, *Is Less More? Towards Better Commonwealth Performance*, Discussion Paper, Commonwealth Financial Accountability Review, Australian Government, Canberra.
- DoH (Department of Health) 2016, *Evaluation Strategy 2019-2019*.
- 2018, *Co-Design in the PHN Commissioning Context*.
- DPC (NSW) (Department of Premier and Cabinet (New South Wales)) 2016a, *C2016-01-Program Evaluation*, <https://arp.nsw.gov.au/c2016-01-program-evaluation> (accessed 18 May 2020).
- 2016b, *NSW Government Program Evaluation Guidelines*.
- DPMC (Department of the Prime Minister and Cabinet) 2017a, *2017-18 Budget: Strengthening Indigenous Research and Evaluation*, Canberra.
- 2017b, *Remote Housing Review: A Review of the National Partnership Agreement on Remote Indigenous Housing and the Remote Housing Strategy (2008-2018)*, Canberra.
- 2018a, *2017-18 Annual Report*, Canberra.
- 2018b, *Indigenous Advancement Strategy Evaluation Framework*, Canberra.
- 2018c, *Indigenous Evaluation Committee Terms of Reference*, Canberra.
- 2018d, *Understanding Family Perspectives of School Attendance in Remote Communities: Evaluation of the Remote School Attendance Strategy*, Canberra.

-
- 2019a, *Closing the Gap Report 2019*, Canberra.
- 2019b, *Data Integration Partnership for Australia*, <https://www.pmc.gov.au/public-data/data-integration-partnership-australia> (accessed 30 April 2020).
- 2019c, *Data Sharing and Release Legislative Reforms Discussion Paper*, Canberra.
- 2019d, *Delivering for Australians — A world-class Australian Public Service: The Government's APS reform agenda*, Canberra.
- 2019e, *Our Public Service, Our Future. Independent Review of the Australian Public Service*, Canberra.
- 2019f, *Third Year Evaluation of the Indigenous Procurement Policy*, Canberra.
- DPMC (NZ) (Department of the Prime Minister and Cabinet (NZ)) 2019, *Terms of Reference - Prime Minister's Science Advisory Committee*.
- Dreise, T. and Mazurski, E. 2018, *Knowledge Exchange, Co-Design and Community-Based Participatory Research and Evaluation in Aboriginal Communities: Literature Review, Case Study and Practical Tips*, NSW Government, Sydney.
- DSS (Department of Social Services) 2015, *Home Interaction Program for Parents and Youngsters (HIPPY): Programme Guidelines 2015-2018*, Canberra.
- 2018, *Department of Social Services Evaluation Policy*, Canberra.
- 2019a, *Data Exchange Policy*, <https://dex.dss.gov.au/policy> (accessed 12 May 2020).
- 2019b, *Footprints in Time: The Longitudinal Study of Indigenous Children - Data User Guide, Release 10.0.*, Canberra.
- 2019c, *Overview of Footprints in Time - The Longitudinal Study of Indigenous Children (LSIC)*, <https://www.dss.gov.au/about-the-department/national-centre-for-longitudinal-studies/overview-of-footprints-in-time-the-longitudinal-study-of-indigenous-children-lsic> (accessed 30 April 2020).
- 2020, *Cashless Debit Card*, Canberra.
- DTF (VIC) (Department of Treasury and Finance (Victoria)) 2019, *The Resource Management Framework for Victorian Government departments*, Government of Victoria, Melbourne, p. 169.
- Eades, S.J., Taylor, B., Bailey, S., Williamson, A.B., Craig, J.C. and Redman, S. 2010, 'The health of urban Aboriginal people: insufficient data to close the gap', *The Medical Journal of Australia*, vol. 193, no. 9, pp. 521–524.
- Edwards, A., Zweigenthal, V. and Olivier, J. 2019, 'Evidence map of knowledge translation strategies, outcomes, facilitators and barriers in African health systems', *Health Research Policy and Systems*, vol. 17, no. 16.
- Effective Change 2016, *Evaluation of Aboriginal Health Case Management and Care Coordination Models in Victoria*, Melbourne.
- Employment and Social Development Canada 2019, *Apprenticeship Grants Evaluation*.
-

-
- Empowered Communities 2015, *Empowered Communities: Empowered Peoples - Design Report*, Wunan Foundation, Sydney.
- EvalPartners 2020, *EvalIndigenous*, <https://www.evalpartners.org/evalindigenous/about> (accessed 1 May 2020).
- Ewen, S., Ryan, T. and Platania-Phung, C. 2019, *Further Strengthening Research Capabilities: A Review and Analysis of the Aboriginal and Torres Strait Islander Health Researcher Workforce*, The Lowitja Institute, Melbourne.
- Executive Office of the President (US) 2019a, *Memorandum for Heads of Executive Departments and Agencies - Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance*, Washington D.C.
- 2019b, *Preparation, Submission, and Execution of the Budget*, Circular No. A-11, Washington D.C.
- Farrington, D.P. 2003, 'Methodological quality standards for evaluation research', *Annals of the American Academy of Political and Social Science*, vol. 587, pp. 49–68.
- Feeney, L., Bauman, J., Chabrier, J., Mehra, G. and Woodford, M. 2015, *Using Administrative Data for Randomized Evaluations*, Abdul Latif Jameel Poverty Action Lab - North America, Cambridge, MA.
- Fetterman, D. 2005, 'Chapter 3: Assessing levels of commitment', *Empowerment Evaluation Principles in Practice*, Guilford Press, New York.
- Fluehr-Lobban, C. (ed) 1991, *Ethics and the Profession of Anthropology: Dialogue for a New Era*, University of Pennsylvania Press, Philadelphia.
- Fogarty, W., Bulloch, H., McDonnell, S. and Davis, M. 2018, *Deficit Discourse and Indigenous Health: How narrative framings of Aboriginal and Torres Strait Islander people are reproduced in policy*, The Lowitja Institute, Melbourne.
- Frydenberg, J. 2018, *First full-time Indigenous Policy Evaluation Commissioner of the Productivity Commission*, Josh Frydenberg MP, <https://joshfrydenberg.com.au/latest-news/first-full-time-indigenous-policy-evaluation-commissioner-of-the-productivity-commission/> (accessed 24 May 2020).
- 2019, *Indigenous Evaluation Strategy: Letter of Direction*.
- Gaetjens, P. and Woolcott, P. 2020, *Open Letter to the Australian Public Service – Thursday 9 April 2020*, Australian Government, Canberra.
- Gamble, J. 2008, *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation, Montreal.
- Geia, L.K., Hayes, B. and Usher, K. 2013, 'Yarning/Aboriginal storytelling: Towards an understanding of an Indigenous perspective and its implications for research practice', *Contemporary Nurse*, vol. 46, no. 1, pp. 13–17.
- Gibb, B., Babyack, S., Stephens, D., Kelleher, K., Hoger, D., Vale, C. and Peersman, G. 2019, *Putting Ethical Principles into Practice — a Protocol to Support Ethical*

Evaluation Practice in Aboriginal and Torres Strait Islander Settings, Working Document-Version 1, BetterEvaluation, Melbourne.

- Gilbert, R., Salanti, G., Harden, M. and See, S. 2005, 'Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002', *International Journal of Epidemiology*, vol. 2005, no. 34, pp. 874–887.
- Goldstein, R. and Hiscox, M.J. 2018, *School Enrolment and Attendance Measure Randomized Controlled Trial: Full Report*, National Indigenous Australians Agency, Canberra, <https://www.niaa.gov.au/resource-centre/indigenous-affairs/school-enrolment-and-attendance-measure-randomized-controlled-trial-full-report> (accessed 12 November 2019).
- Government of Canada 2016a, *Directive on Results*, <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31306> (accessed 23 May 2018).
- 2016b, *Policy on Results*, <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31300> (accessed 24 April 2020).
- 2019, *Audit and evaluation reports*, <https://www.canada.ca/en/treasury-board-secretariat/corporate/audit-evaluation-reports.html> (accessed 28 April 2020).
- 2020a, *Crown-Indigenous Relations and Northern Affairs Canada*, <https://www.canada.ca/en/crown-indigenous-relations-northern-affairs.html> (accessed 28 April 2020).
- 2020b, *Evaluation learning series 2019-2020*, <https://www.cspsefpc.gc.ca/events/eval-serie2019-20/index-eng.aspx> (accessed 27 April 2020).
- 2020c, *Indigenous and Northern Affairs Canada*, <https://www.canada.ca/en/indigenous-northern-affairs.html> (accessed 28 April 2020).
- 2020d, *Indigenous Services Canada*, <https://www.canada.ca/en/indigenous-services-canada.html> (accessed 28 April 2020).
- 2020e, *Treasury Board of Canada Secretariat*, <https://www.canada.ca/en/treasury-board-secretariat.html> (accessed 29 April 2020).
- Government of the United Kingdom 2019, *What works network*, <https://www.gov.uk/guidance/what-works-network#the-what-works-network> (accessed 28 April 2020).
- Government Social Research (UK) 2019, *Government Social Research Technical Framework*.
- Gray, M. and Bray, J.R. 2019, *Evaluation in the Australian Public Service: Current State of Play, Some Issues and Future Directions*, Australia and New Zealand School of Government, Canberra.
- Greene, J.C., Caracelli, V.J. and Graham, W.F. 1989, 'Toward a conceptual framework for mixed-method evaluation designs', *Educational Evaluation and Policy Analysis*, vol. 11, no. 3, pp. 255–274.

-
- Grey, K., Yamaguchi, J., Williams, E., Davis, V., Foster, D., Gibson, J. and Dunnett, D. 2018, 'The Strength of Indigenous Australian Evaluators and Indigenous Evaluation: A SnapShot of "Ways of Knowing and Doing" Reflecting on the 2014 Darwin Conference of the Australasia Evaluation Society', *Indigenous Evaluation: New Directions for Evaluation*, vol. 159, pp. 79–95.
- Gruen, N. 2016a, *Why Australia needs an evaluator-general*, <https://www.themandarin.com.au/64566-nicholas-gruen-evaluator-general-part-two/> (accessed 2 May 2020).
- 2016b, *Why we accept travesties of 'evidence-based' policymaking*, <https://www.themandarin.com.au/64557-nicholas-gruen-evidence-based-policy-part-one/> (accessed 14 May 2020).
- 2018, *Submission to the Independent Review of the Australian Public Service*, Melbourne.
- GSR (Government Social Research Profession) nd, *About us*, <https://www.gov.uk/government/organisations/civil-service-government-social-research-profession/about> (accessed 29 April 2020).
- Halligan, J. 2003, 'Public-sector reform and evaluation in Australia and New Zealand', in Wollmann, H. (ed), *Evaluation in Public-sector Reform: Concepts and Practice in International Perspective*, Edward Elgar Publishing, Cheltenham, United Kingdom, pp. 80–103.
- Hawke, L. and Wanna, J. 2010, 'Australia after budgetary reform: a lapsed pioneer or decorative architect', in Wanna, J., Jensen, L. and de Vries, J. (eds), *The Reality of Budgetary Reform in OECD Nations: Trajectories and Consequences*, Edward Elgar Publishing, Cheltenham, United Kingdom, pp. 65–90.
- Haynes, A., Rowbotham, S., Redman, S. and Brennan, S. 2018, 'What can we learn from interventions that aim to increase policymakers' capacity to use research? A realist scoping review', *Health Research Policy and Systems*, vol. 16, no. 31, pp. 1–27.
- Haynes, L., Service, O., Goldacre, B. and Torgerson, D. 2012, *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*, Behavioural Insights Team, Cabinet Office, United Kingdom.
- Hayward, B. and Villa, L. 2015, *'My Home is My Marae' — Kaupapa Māori Evaluation: Community Report*.
- Hayward, R.J., Kay, J., Lee, A., Page, E., Patel, N., Santos, J.P., Perry, L., Valyraki, A. and Payne, H. 2014, 'Evaluation under contract: government pressure and the production of policy research', *Public Administration*, vol. 92, no. 1, pp. 224–239.
- Healing Foundation 2017, *Collaboration and co-design when evaluating intergenerational trauma projects*, Text, Child Family Community Australia (Australian Institute of Family Studies), <https://aifs.gov.au/cfca/2017/05/09/collaboration-and-co-design-when-evaluating-intergenerational-trauma-projects> (accessed 22 May 2020).

-
- Healthhabitat 2020, *How Might Housing for Health Projects Improve Health? — Housing for Health — The Guide*, <http://www.housingforhealth.com/about/the-health-story/4-how-might-housing-for-health-projects-improve-health/> (accessed 3 March 2020).
- Helme, S. and Lamb, S. 2011, *Closing the School Completion Gap for Indigenous Students*, Resource Sheet 6, Closing the Gap Clearinghouse.
- HM Treasury nd, *Government Analytical Evaluation Capabilities Framework*, London.
- nd, *Welcome to the Government Analytical Evaluation Capabilities Framework Interactive Tool*, London.
- 2011, *The Magenta Book: Guidance for Evaluation*, London.
- 2012, *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence (Supplementary Magenta Book Guidance)*, London.
- 2015, *The Aqua Book: Guidance on Producing Quality Analysis for Government*, London.
- 2018, *The Green Book: Appraisal and Evaluation in Central Government*, London.
- 2020a, *Magenta Book 2020 — Supplementary Guide: Handling Complexity in Policy Evaluation*, London.
- 2020b, *Magenta Book 2020 — Supplementary Guide: Realist Evaluation*, London.
- 2020c, *The Magenta Book: Central Government Guidance on Evaluation*, London.
- House of Representatives Standing Committee on Aboriginal Affairs 1979, *Aboriginal Health*, Canberra.
- House of Representatives Standing Committee on Finance and Public Administration 1990, *Not Dollars Alone: Review of the Financial Management Improvement Program*.
- Hudson, M., Milne, M., Reynolds, P., Russell, K. and Smith, B. 2010, *Te Ara Tika: Guidelines for Māori Research Ethics - A Framework for Researchers and Ethics Committee Members*, Health Research Council of New Zealand, Auckland.
- Hudson, S. 2016, *Mapping the Indigenous Program and Funding Maze*, Research Report no. 18, Centre for Independent Studies, Sydney.
- 2017, *Evaluating Indigenous Programs: a Toolkit for Change*, Research Report no. 28, Centre for Independent Studies, Sydney.
- 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.
- 2016, *Let's Talk About Success: Exploring Factors Behind Positive Change in Aboriginal Communities*, Working Paper 109, Centre for Aboriginal Economic Policy Research, Canberra.
- 2017, *The Cashless Debit Card Trial Evaluation: A Short Review*, CAEPR Topical Issues 1, Centre for Aboriginal Economic Policy Research, Canberra.

-
- IHS (Indian Health Service) nd, *Evaluation policy*, <https://www.ihs.gov/dper/evaluation/evaluation-policy/> (accessed 29 April 2020).
- Independent Māori Statutory Board 2017, *Schedule of Issues of Significance to Māori in Tāmaki Makaurau and Māori Plan 2017*, Auckland, New Zealand.
- nda, *About us*, <https://www.imsb.maori.nz/what-we-do/issues-of-significance/> (accessed 21 April 2020a).
- ndb, *What we do*, <https://www.imsb.maori.nz/what-we-do/issues-of-significance/> (accessed 21 April 2020b).
- Indigenous Services Canada 2018, *Evaluation of the On-Reserve Income Assistance Program*, Evaluation, Performance Measurement and Review Branch, Canada.
- 2019, *Evaluation reports*, <https://www.sac-isc.gc.ca/eng/1550775205150/1550775238874> (accessed 28 April 2020).
- 2020, *Results at a Glance - Evaluation of the First Nations and Inuit Home and Community Care Program*, <https://www.sac-isc.gc.ca/eng/1568398721350/1568398756626> (accessed 21 April 2020).
- Inside Policy 2018, *An Evaluation of the Family Advocacy and Support Services: Final Report, prepared by Inside Policy for the Australian Government Attorney-General's Department*, Attorney-General's Department, Canberra.
- IRDC (International Development Research Centre) 2011, 'In Conversation: Michael Quinn Patton', <https://www.idrc.ca/en/research-in-action/conversation-michael-quinn-patton> (accessed 22 May 2020).
- Jacob, S. and Desautels, G. 2014, 'Assessing the quality of Aboriginal program evaluations', *Canadian Journal of Program Evaluation*, vol. 29, no. 1, pp. 62–86.
- and — 2013, 'Evaluation of Aboriginal programs: What place is given to participation and cultural sensitivity?', *The International Indigenous Policy Journal*, vol. 4, no. 2, pp. 1–29.
- James, M. 2013, 'Designing evaluation strategies', *Better Indigenous Policies: The Role of Evaluation*, Productivity Commission, Canberra, pp. 107–118.
- JCOCTG (Joint Council on Closing the Gap) 2019a, *Communique — First Meeting of the Joint Council on Closing The Gap*, Brisbane.
- 2019b, *Communique — Second Meeting of the Joint Council on Closing The Gap*, Adelaide.
- 2019c, *Partnership Agreement on Closing the Gap 2019-2029*, Canberra.
- Jeffares, B., Boston, J., Gerrard, J., Hendy, S. and Larner, W. 2019, 'Science advice in New Zealand: opportunities for development', *Policy Quarterly*, vol. 15, no. 2, pp. 62–71.
- Jones, R., Thurber, K.A., Chapman, J., D'Este, C., Dunbar, T., Wenitong, M., Eades, S.J., Strelein, L., Davey, M., Du, W., Olsen, A., Smylie, J.K., Banks, E. and Lovett, R. 2018, 'Study protocol: Our Cultures Count, the Mayi Kuwayu Study, a national longitudinal

-
- study of Aboriginal and Torres Strait Islander wellbeing', *BMJ Open*, vol. 8, no. 6, p. e023861.
- Kelagher, M., Luke, J., Ferdinand, A., Chamravi, D., Ewen, S. and Paradies, Y. 2018a, *An Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health*, The Lowitja Institute, Melbourne.
- , ——, ——, ——, —— and —— 2018b, *An Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health Users Guide*, The Lowitja Institute, Melbourne.
- KPMG 2018, *Maranguka Justice Reinvestment Project Impact Assessment*, Maranguka and Just Reinvest NSW, Sydney.
- 2019, *Final Report for the Evaluation of the Individual Placement and Support Trial*, Department of Social Services, Canberra.
- Laycock, A., Walker, D., Harrison, N. and Brands, J. 2011, *Researching Indigenous Health: A Practical Guide for Researchers*, The Lowitja Institute, Melbourne.
- Leeson, S., Smith, C. and Rynne, J. 2016, 'Yarning and appreciative inquiry: The use of culturally appropriate and respectful research methods when working with Aboriginal and Torres Strait Islander women in Australian prisons', *Methodological Innovations*, vol. 9, pp. 1–17.
- Legal Aid WA 2019, *Blurred Borders Process Review and Evaluation Summary Report*, Perth.
- Leonhardt, D. 2014, 'The Quiet Movement to Make Government Fail Less Often', *The New York Times*, 15 July.
- Li, J.-L. 2017, 'Cultural barriers lead to inequitable healthcare access for Aboriginal Australians and Torres Strait Islanders', *Chinese Nursing Research*, vol. 4, no. 4, pp. 207–210.
- Liamputtong, P. 2013, *Qualitative Research Methods*, Oxford University Press, Victoria.
- Library and Archives Canada 2012, *Terminology Guide: Research on Aboriginal Heritage*.
- Liddell, M., Barnett, T., McEachran, J. and Roost, F.D. 2010, *HIPPY national evaluation: Interim report*, Brotherhood of St Laurence, Melbourne.
- , ——, Roost, F.D. and McEachran, J. 2011, *Investing in our Future: An evaluation of the rollout of the Home Interaction Program for Parents and Youngsters (HIPPY)*, Brotherhood of St Laurence, Melbourne.
- Lovett, R. 2016, 'Aboriginal and Torres Strait Islander community wellbeing: Identified needs for statistical capacity', *Indigenous Data Sovereignty: Toward an Agenda*, Research Monograph No. 39, Centre for Aboriginal Economic Policy Research, Canberra.
- Mā te Rae 2015, *Mā te Rae — Māori Evaluation Association (Facebook Page)*, https://www.facebook.com/pg/MaTeRae/about/?ref=page_internal (accessed 1 May 2020).

-
- nd, *Indigenous peoples' conference on evaluation*, <https://www.maaterae.co.nz/> (accessed 5 May 2020).
- Mackay, K. 1998, *The Development of Australia's Evaluation System*, ECD 4, ECD Working Paper Series, The World Bank, Washington, D.C.
- 2011a, *The Australian Government's Performance Framework*, Evaluation Capacity Development Working Paper, ECD 25, The World Bank, Washington, D.C.
- 2011b, 'The Performance Framework of the Australian Government, 1987 to 2011', presented at 7th Annual Meeting on Performance and Results, Paris, 9 November.
- Maiaam Nayri Wingara Indigenous Data Sovereignty Collective and AIGI (Maiaam Nayri Wingara Indigenous Data Sovereignty Collective and Australian Indigenous Governance Institute) 2018, *Indigenous Data Sovereignty: Communique*, Indigenous Data Sovereignty Summit, Canberra.
- Maloney, J. 2017, 'Evaluation: what's the use?', *Evaluation Journal of Australasia*, vol. 17, no. 4, pp. 25–38.
- Marcel, M. 2015, *World Bank Blogs*, <https://blogs.worldbank.org/governance/value-evaluations-asking-right-questions> (accessed 15 September 2019).
- Martin, K. 2003, 'Ways of knowing, being and doing: A theoretical framework and methods for Indigenous and Indigenist re-search', *Voicing Dissent*, University of Queensland Press, Brisbane, pp. 203–257.
- Mavromaras, K., Moskos, M., Mahuteau, S., Isherwood, L., Goode, A., Walton, H., Smith, L., Wei, Z. and Flavel, J. 2018, *Evaluation of the NDIS: Final Report*, National Institute of Labour Studies, Flinders University, Adelaide.
- Mayi Kuwayu 2019, "*Can I access the data?*", <https://mkstudy.com.au/can-i-access-the-data/> (accessed 5 November 2020).
- Mayne, J. 2010, 'Building an evaluative culture: the key to effective evaluation and results management', *The Canadian Journal of Program Evaluation*, vol. 24, no. 2, pp. 1–30.
- McDonald, C. 2009, *Modern Budget Performance and the Impact on Parliaments: The Australian Experience*, Senate Finance and Public Administration Committee, Canberra.
- McPhee, I. 2011, *The Evolving Role and Mandate of the ANAO Since Federation*, Senate Occasional Lecture, Australian National Audit Office, Canberra.
- Meit, M., Hafford, C., Fromknecht, C., Knudson, A., Gilbert, T. and Noelle, M. 2016, *Tribal Health Profession Opportunity Grants (HPOG) Program Evaluation*, OPRE Report 2016-38, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- Mertens, D.M. and Wilson, A.T. 2012, 'Chapter 9: Evaluation Designs', *Program Evaluation Theory and Practice: A Comprehensive Guide*, Guilford Press, New York.
- Milligan, V., Phillips, R., Easthope, H., Liu, E. and Memmott, P. 2011, *Urban Social Housing for Aboriginal People and Torres Strait Islanders: Respecting Culture and Adapting Services*, Australian Housing and Urban Research Institute.

-
- Milroy, A., Nelson-Kelly, S., Hennessey, K. and Morgan, J. 2019, *Indigenous Values for the APS*, Submission to the APS Review, Australia and New Zealand School of Government.
- Minister for Māori Development 2019, *Positioning Te Puni Kōkiri for Strategic Impact*.
- Ministry of Justice (NZ) 2020, *Treaty of Waitangi*, <https://www.justice.govt.nz/about/learn-about-the-justice-system/how-the-justice-system-works/the-basis-for-all-law/treaty-of-waitangi/> (accessed 28 January 2020).
- Mitchell, E., Bandara, P. and Smith, V. 2018, *Tackling Indigenous Smoking Program – Final Evaluation Report*, prepared for the Australian Government Department of Health, CIRCA, Leichhardt, NSW.
- Mitton, C., Adair, C.E., McKenzie, E., Patten, S.B. and Perry, B.W. 2007, 'Knowledge Transfer and Exchange: Review and Synthesis of the Literature', *The Milbank Quarterly*, vol. 85, no. 4, pp. 729–768.
- Moewaka Barnes, H. 2000, 'Kaupapa Māori: explaining the ordinary', *Pacific Health Dialogue*, vol. 7, no. 2, pp. 13–16.
- 2013, 'Better Indigenous policies: an Aotearoa New Zealand perspective on the role of evaluation', *Better Indigenous Policies: The Role of Evaluation*, Productivity Commission, Canberra, pp. 155–182.
- Moran, M. 2017, *Submission in Response to the Queensland Productivity Commission Draft Report*, Queensland Productivity Commission.
- Moreton-Robinson, A. 2004, 'Whiteness, epistemology and Indigenous representation', *Whitening Race: Essays in Social and Cultural Criticism*, Aboriginal Studies Press, Canberra, pp. 75–78.
- Morrison, S. 2020, *Communique - Coalition of Aboriginal and Torres Strait Islander Peak Organisations*, <https://www.pm.gov.au/media/communique-coalition-aboriginal-and-torres-strait-islander-peak-organisations> (accessed 10 February 2020).
- Morton, D. and Cook, B. 2015, *Evaluators and the Enhanced Commonwealth Performance Framework*, Department of Finance.
- Muir, S. and Dean, M. 2017, *Evaluating the Outcomes of Programs for Indigenous Families and Communities*, CFCA Practice Resource.
- Nakata, M. 2002, 'Indigenous knowledge and the cultural interface: underlying issues at the intersection of knowledge and information systems', *IFLA Journal*, vol. 28, pp. 281–291.
- NATSILMH (National Aboriginal and Torres Strait Islander Leadership in Mental Health) 2015, *Gayaa Dhuwi Declaration*.
- NCIRS (National Centre for Immunisation Research and Surveillance) 2014, *Evaluation of the National Human Papillomavirus Vaccination Program*, Department of Health, Canberra.
- NCLD (National Centre for Longitudinal Data) 2018, *What is meant by cultural competency and cultural standpoint?*, https://ada.edu.au/wp-content/uploads/2018/11/What_is_meant_by_cultural_competency_and_cultural_standpoint.pdf (accessed 30 April 2020).

-
- NCS (National Curriculum Services) 2012, *What Works. The Work Program. Success in remote schools: a Research Study of Eleven Improving Remote Schools*, Melbourne.
- NHMRC, ARC and UA (National Health and Medical Research Council, Australian Research Council and Universities Australia) 2018, *National Statement on Ethical Conduct in Human Research*, National Health and Medical Research Council, Canberra.
- NHMRC (National Health and Medical Research Council) 2014, *Ethical Considerations in Quality Assurance and Evaluation Activities*, National Health and Medical Research Council, Canberra.
- 2018a, *Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities: Guidelines for Researchers and Stakeholders*, National Health and Medical Research Council, Canberra.
- 2018b, *Keeping research on track II*, Commonwealth of Australia, Canberra.
- 2020, *Ethical guidelines for research with Aboriginal and Torres Strait Islander Peoples*, <https://www.nhmrc.gov.au/research-policy/ethics/ethical-guidelines-research-aboriginal-and-torres-strait-islander-peoples> (accessed 4 February 2020).
- NIAA (National Indigenous Australians Agency) 2019, *2019-20 Annual Evaluation Work Plan*, Canberra.
- 2020, *Funding Under the IAS*, <https://www.niaa.gov.au/indigenous-affairs/grants-and-funding/funding-under-ias> (accessed 4 May 2020).
- nd, *Empowered Communities*, <https://www.niaa.gov.au/indigenous-affairs/empowered-communities> (accessed 3 April 2020).
- Nicholls, J., Lawlor, E., Neitzert, E. and Goodspeed, T. 2012, *A Guide to Social Return on Investment*, The Social Return on Investment Network, United Kingdom, <http://www.socialvalueuk.org/app/uploads/2016/03/The%20Guide%20to%20Social%20Return%20on%20Investment%202015.pdf> (accessed 2 February 2020).
- Nickerson, R. 1998, 'Confirmation bias: a ubiquitous phenomenon in many guises', *Review of General Psychology*, vol. 2, no. 2, pp. 175–220.
- Nous Consulting 2019, *Nous Review of IUIH — Summary Report 2019*, <https://www.iuih.org.au/Blog/ArticleID/7056/Nous-Review-of-IUIH-Summary-Report-2019> (accessed 21 April 2020).
- NSW Treasury 2018, *Program Evaluation*, Treasury Circular, TC 18-03, NSW Government, Sydney.
- NT Government 2019, *Submission by the Northern Territory Government to the Productivity Commission's Draft Report into expenditure in the area of children and family services in the Northern Territory*.
- OECD (Organisation for Economic Co-Operation and Development) 1987, *The Control and Management of Government Expenditure*, Paris.
- 2017, *Fostering Innovation in the Public Sector*, OECD Publishing, Paris.

-
- 2019, *Linking Indigenous Communities with Regional Development*, OECD Rural Policy Reviews, Paris.
- 2020a (Forthcoming), *Institutionalisation, Quality and Use of Policy Evaluation: Lessons from Countries Experience*, Paris.
- 2020b, *Linking Indigenous Communities with Regional Development in Canada*, OECD Rural Policy Reviews, Paris.
- Office of Evaluation Sciences (US) 2020, *Evidence Act Toolkits*, <https://oes.gsa.gov/toolkits/> (accessed 27 April 2020).
- nd, *Mission*, <https://oes.gsa.gov/about/> (accessed 27 April 2020).
- Office of the Prime Minister's Chief Science Advisor (NZ) 2018, *Chief Science Advisor Forum*.
- O'Hara, J., McGill, B., Ahmed, N., McElduff, C. and Rissel, C. 2016, *The Get Healthy Information and Coaching Service Aboriginal Program: Evidence and Evaluation Report*, NSW Ministry of Health and the Apple Institute, Sydney.
- OIPC (Office of Indigenous Policy Coordination) 2006, *Office of Indigenous Policy Coordination Evaluation Plan for Whole-of-Government Activities in Indigenous Affairs 2006-09*, Canberra.
- 2010, *Directory of Commonwealth Government evaluations, audits and reviews of Indigenous-specific and relevant mainstream programs 2002-2014*, Canberra.
- OMB (Office of Management and Budget) 2019, *A Budget for a Better America: Fiscal Year 2020 Analytical Perspectives*, U.S. Government, Washington, D.C.
- Ombudsman New South Wales 2019, *OCHRE Review Report*, Sydney.
- ONDC (Office of the National Data Commissioner) 2019, *About us*, <https://www.datacommissioner.gov.au/about/about-us> (accessed 12 May 2020).
- OPRE (Office of Planning, Research and Evaluation) 2020, *What we do*, <https://www.acf.hhs.gov/opre/about/what-we-do> (accessed 11 May 2020).
- ORIMA Research 2017, *Cashless Debit Card Trial Evaluation - Final Evaluation Report*, Department of Social Services, Canberra.
- Osborne, K., Baum, F. and Brown, L. 2013, *What Works? A Review of Actions Addressing the Social and Economic Determinants of Indigenous Health*, Issues Paper 7, Closing the Gap Clearinghouse.
- Osrin, D., Azad, K., Fernandez, A., Manandhar, D.S., Mwansambo, C.W., Tripathy, P. and Costello, A.M. 2009, 'Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia', *Bulletin of the World Health Organization*, vol. 87, pp. 772–779.
- Pagan, B. 2016, 'Treasury Board Policy on Results: Financial Management Institute - PD Week'.

-
- Patton, M.Q. 1994, 'Developmental evaluation', *Evaluation Practice*, vol. 15, no. 3, pp. 311–319.
- 2013a, 'The UFE Checklist', in Ramirez, R. and Brodhead, D. (eds), *Utilization Focused Evaluation: A Primer for Evaluators*, Southbound, Penang, Malaysia, pp. 89–102.
- 2013b, *Utilization-Focused Evaluation Checklist*, Evaluation Checklists, The Evaluation Centre, Western Michigan University, Kalamazoo, Michigan.
- 2016, 'What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models', *American Journal of Evaluation*, vol. 37, no. 2, pp. 250–265.
- , McKegg, K. and Wehipeihana, N. 2015, *Developmental evaluation exemplars: Principles in practice*, The Guilford Press, New York/ London.
- Pawson, R. and Tilley, N. 2004, *Realist Evaluation*, British Cabinet Office, United Kingdom.
- Payne, M., Chi-Chung, C. and Knight, P. 2019, 'A model for data sharing and research: the Indigenous Research Exchange', presented at AIATSIS National Indigenous Research Conference, Brisbane.
- PC 2010, *Strengthening Evidence-based Policy in the Australian Federation, Volume 2: Background Paper*, Productivity Commission, Canberra.
- 2013, *Better Indigenous Policies: the Role of Evaluation*, Roundtable Proceedings, Productivity Commission, Canberra.
- 2015, *National Indigenous Reform Assessment Agreement*, Performance Assessment, 2013–14, Canberra.
- 2016a, *Indigenous Primary School Achievement*, Commission Research Paper, Canberra.
- 2016b, *National Education Evidence Base*, Report no. 80, Canberra.
- 2017a, *Data Availability and Use*, Report no. 82, Canberra.
- 2017b, *Introducing Competition and Informed User Choice into Human Services: Reforms to Human Services*, Report no. 85, Canberra.
- 2019a, *A Better Way to Support Veterans*, Report no. 93, Canberra.
- 2019b, *Mental Health - Draft Report*, Canberra.
- 2020, *Expenditure on Children in the Northern Territory*, Study Report, Canberra.
- Peiris, D., Wright, L., Corcoran, K., News, M. and Turnbull, F. 2017, *1 Deadly Step: Process evaluation of a chronic disease screening program in NSW Aboriginal communities*, The George Institute for Global Health, Newcastle.
- Perrin, B. 2012, *Linking Monitoring and Evaluation to Impact Evaluation*, No. 2, Impact Evaluation Notes, InterAction, Washington, D.C.

-
- Pleger, L. and Sager, F. 2018, 'Betterment, undermining, support and distortion: A heuristic model for the analysis of pressure on evaluators', *Evaluation and Program Planning*, vol. 69, pp. 166–172.
- Plowman, C. 2008, 'Evaluating Indigenous Programs — Creating Public Value?'
- PM&C (Department of the Prime Minister and Cabinet) 2018, *2018-19 Evaluation Work Plan*, Canberra.
- Podger, A. 2018, 'Making "accountability for results" really work?', in Su, T.-T., Wanna, J., Chan, H.S., Niu, M. and Niu, M. (eds), *Value For Money: Budget and Financial Management Reform in the People's Republic of China, Taiwan and Australia*, ANU Press, Acton.
- Posavac, E.J. and Carey, R.G. 1997, *Program Evaluation Methods and Case Studies*, 5th edn, Prentice-Hall, New Jersey.
- Power, A., Woodrow, C., Orlando, J. and Western Sydney University. Centre for Educational Research 2016, *Evaluation Report Early Childhood Schools and Koori Pre-Program*, University of Western Sydney, Penrith.
- Priest, N., Mackean, T., Waters, E., Davis, E. and Riggs, E. 2009, 'Indigenous Child Health Research: a Critical Analysis of Australian Studies', *Australia and New Zealand Journal of Public Health*, vol. 33, pp. 55–63.
- Purdie, N. and Buckley, S. 2010, *School Attendance and Retention of Indigenous Australian Students*, Issues Paper 1, Closing The Gap Clearinghouse.
- PWC Indigenous Consulting 2018, *Review of the Northern Territory Sobering Up Shelters*, Darwin.
- QPC (Queensland Productivity Commission) 2017, *Service Delivery in Remote and Discrete Aboriginal and Torres Strait Islander Communities*, Brisbane.
- Quinn, E., O'Hara, B.J., Ahmed, N., Winch, S., McGill, B., Banovic, D., Maxwell, M. and Rissel, C. 2017, 'Enhancing the get healthy information and coaching service for aboriginal adults: evaluation of the process and impact of the program', *International Journal for Equity in Health*, vol. 16, no. 168, pp. 1–11.
- RANZCP (The Royal Australian and New Zealand College of Psychiatrists) 2020, *The Dance of Life*, <https://www.ranzcp.org/practice-education/indigenous-mental-health/aboriginal-torres-strait-islander-mental-health/the-dance-of-life> (accessed 11 February 2020).
- RCAGA (Royal Commission on Australian Government Administration) 1976, *Royal Commission on Australian Government Administration: Report*, Australian Government Publishing Service, Canberra.
- Research Data Alliance International Indigenous Data Sovereignty Group 2019, *CARE Principles for Indigenous Data Governance*, Global Indigenous Data Alliance.
- Restorative Justice Evaluation Team 2018, *Restorative Justice Project 12-Month Program Evaluation*, Department of Child Safety, Youth and Women, Brisbane.

-
- Rigney, L. 2006, 'Indigenous Australian Views on Knowledge Production and Indigenist Research', *Indigenous Peoples' Wisdom and Power: Affirming Our Knowledge Through Narratives*, Hampshire and Burlington, Ashgate, pp. 32–49.
- Rocco, T.S. 2010, 'Criteria for evaluating qualitative studies', *Human Resource Development International*, vol. 13, no. 4, pp. 375–378.
- Rogers, P. 2009a, 'Learning from the evidence about evidence-based policy', *Strengthening evidence-based policy in the Australian Federation: Roundtable proceedings*, vol 1, Productivity Commission, Canberra.
- 2009b, 'Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation', *Journal of Development Effectiveness*, vol. 1, no. 3, pp. 217–226.
- Rossi, P., Lipsey, M. and Henry, G. 2018, *Evaluation: A systematic approach*, Ed. 8, Sage Publications, United States.
- Russell-Mundine, G. 2017, 'The Cultural Capability of New South Wales Public Servants', *Transforming the Relationship Between Aboriginal Peoples and the NSW Government: Aboriginal Affairs NSW Research Agenda 2018-2023*, Aboriginal Affairs NSW, Sydney, pp. 43–52.
- Ryan, B. 2003, "“Death by evaluation”?: Reflections on monitoring and evaluation in Australia and New Zealand", *Evaluation Journal of Australasia*, vol. 3, no. 1, pp. 6–16.
- SAHMRI (South Australian Health and Medical Research Institute) 2016, *Evaluation of the IUIH CCSS program Final report*, Adelaide.
- Schoenefeld, J. and Jordan, A. 2017, 'Governing policy evaluation? Towards a new typology', *Evaluation*, vol. 23, no. 3, pp. 274–293.
- SCRGSP (Steering Committee for the Review of Government Service Provision) 2016, *Overcoming Indigenous Disadvantage: Key Indicators 2016*, Productivity Commission, Canberra.
- 2017, *Indigenous Expenditure Report 2017*, Productivity Commission, Canberra.
- 2019, *Report on Government Services 2019*, Productivity Commission, Canberra.
- Services Australia 2019, *ABSTUDY — Who can get it — Services Australia*, <https://www.servicesaustralia.gov.au/individuals/services/centrelink/abstudy/who-can-get-it> (accessed 12 May 2020).
- Shand, D. 1998, 'Evaluation Capacity and the Public Sector', in Mackay, K. (ed), *Public Sector Performance — The Critical Role of Evaluation: Selected proceedings from a World Bank Seminar*, The World Bank, Washington, D.C., pp. 15–18.
- Sharp, C.A. 2003, 'Development of program evaluation in Australian and the Australasian Evaluation Society - the early decades', *Evaluation Journal of Australasia*, vol. 3, no. 2, pp. 6–16.
- Shergold, P. 2015, *Learning from Failure: Why large government policy initiatives have gone so badly wrong in the past and how the chances of success in the future can be improved*, Australian Public Service Commission, Canberra.

-
- Smith, J. and Adamson, E. 2018, *Process Evaluation of the Banned Drinker Register in the Northern Territory*, Menzies School of Health Research, Darwin.
- , Pollard, K., Robertson, K. and Shalley, F. 2018, *Strengthening Evaluation in Indigenous Higher Education Contexts in Australia: Equity Fellowship Report*, National Centre for Student Equity in Higher Education, Perth.
- Smith, L. 2018, 'Turning research into action through knowledge translation', *Health Voices*, vol. 2020, no. 1.
- Smith, L.T. 1999, *Decolonizing Methodologies: Research and Indigenous Peoples*, Zed Books, London.
- Smyth, C., Gendera, S., Giuntoli, G., Hill, T., Katz, I. and Asante, D. 2017, *Evaluation of Settlement Grants - Final Report, Prepared for Australian Government Department of Social Services*, SPRC Report, 12/17, Social Policy Research Centre, UNSW, Sydney.
- Social Compass 2019, *Evaluation of the Pilot Program of Specialist Domestic Violence Units and Health Justice Partnerships Established under the Women's Safety Package: Final Report, prepared by Social Compass for the Attorney General's Department.*, Attorney-General's Department, Canberra.
- SPEaR (Social Policy Evaluation and Research Committee) 2008, *SPEaR Good Practice Guidelines*, June, New Zealand.
- SSCSW (Senate Standing Committee on Social Welfare) 1979, *Through a Glass, Darkly - Evaluation in Australian Health and Welfare Services*, Canberra.
- TBS (Treasury Board of Canada Secretariat) 2010, *About the Centre of Excellence for Evaluation*, <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/about-centre-excellence-evaluation.html> (accessed 23 May 2018).
- 2018, *Treasury Board of Canada Secretariat mandate*, <https://www.canada.ca/en/treasury-board-secretariat/corporate/mandate.html> (accessed 29 April 2020).
- Te Arawhiti (Office for Māori Crown Relations) 2019a, *Māori Crown Relations Capability Framework for the Public Service*, Wellington.
- 2019b, *Māori Crown Relations Capability Framework for the Public Service - Individual Capability Component*, Wellington.
- 2019c, *Māori Crown Relations Capability Framework for the Public Service - Organisational Capability Component*, Wellington.
- nd, *Crown engagement with Māori*, <https://www.tearawhiti.govt.nz/tools-and-resources/crown-engagement-with-maori/> (accessed 29 April 2020).
- The Lowitja Institute 2020, *Knowledge translation*, <https://www.lowitja.org.au/page/research/knowledge-translation> (accessed 30 April 2020).

-
- Thurber, K.A., Banks, E. and Banwell, C. 2015, 'Cohort profile: Footprints in time, the Australian Longitudinal Study of Indigenous Children', *International Journal of Epidemiology*, vol. 44, no. 3, pp. 789–800.
- TPK (Te Puni Kōkiri (Ministry of Māori Development)) 1999, *Evaluation for Māori - Guidelines for Government Agencies*, New Zealand.
- 2015, *The Whanau Ora Outcomes Framework*, New Zealand.
- 2020, *Roles and functions*, <https://www.tpk.govt.nz/en/mo-te-puni-kokiri/what-we-do/our-role-and-functions> (accessed 29 April 2020).
- Treasury and Trade (QLD) 2014, *Queensland Government Program Evaluation Guidelines*, Queensland Government.
- Tribal Epidemiology Centers 2013, *Best Practices in American Indian & Alaska Native Public Health*.
- 2020, *7 Core Functions of Tribal Epidemiology Centers*, <https://tribalepicenters.org/7-core-functions/> (accessed 15 March 2019).
- Tune, D. 2010, 'Evaluation: Renewed Strategic Emphasis', Presentation to the Canberra Evaluation Forum, Canberra.
- UN (United Nations) 2007, *United Nations Declaration on the Rights of Indigenous Peoples*.
- 2016, *UNEG Norms and Standards for Evaluation*, Evaluation Group, New York.
- UNEG (United Nations Evaluation Group) 2016, *Evaluation Competencies for the UN System for Evaluators, Managers and Commissioners of Evaluation*, United Nations Evaluation Group.
- University of Newcastle 2016, *48 Hour Follow Up Evaluation: Final Report*.
- Victorian Auditor-General 2014, *Accessibility of Mainstream Services for Aboriginal Victorians*, Government of Victoria, Melbourne, p. 70.
- Walter, M. 2016, 'Data politics and Indigenous representation in Australian statistics', in Kukutai, T. and Taylor, J. (eds), *Indigenous data sovereignty: toward an agenda*, Research monograph no. 38, Australian National University Press, Canberra, pp. 79–97.
- and Andersen, C. 2013, *Indigenous statistics: A quantitative research methodology*, Left Coast Press.
- and Suina, M. 2019, 'Indigenous data, indigenous methodologies and indigenous data sovereignty', *International Journal of Social Research Methodology*, vol. 22, no. 3, pp. 233–243.
- Ward, J., McGregor, S., Guy, R.J., Rumbold, A.R., Garton, L., Silver, B.J., Taylor-Thompson, D., Hengel, B., Knox, J., Dyda, A., Law, M.G., Wand, H., Donovan, B., Fairley, C.K., Skov, S., Ah Chee, D., Boffa, J., Glance, D., McDermott, R., Maher, L. and Kaldor, J.M. 2013, 'STI in remote communities: improved and enhanced primary health care (STRIVE) study protocol: a cluster randomised controlled trial comparing "usual practice" sti care to enhanced care in remote primary health care services in

-
- australia', *BMC Infectious Diseases*, vol. 13, no. 425, <https://bmcinfectdis.biomedcentral.com/track/pdf/10.1186/1471-2334-13-425> (accessed 12 October 2019).
- Ware, V.-A. 2013, *Housing Strategies that Improve Indigenous Health Outcomes*, Resource sheet 25, Closing the Gap Clearinghouse.
- Watson, N. and Wooden, M. 2004, 'Sample attrition in the HILDA survey', *Australian Journal of Labour Economics*, vol. 7, no. 2, pp. 293–308.
- Webb, R. 2002, *The Commonwealth Budget: Process and Presentation*, Research Paper, No. 10 2001-02, Department of the Parliamentary Library, Canberra.
- Wehipeihana, N. 2019a, 'Increasing cultural competence in support of Indigenous-led evaluation: a necessary step toward Indigenous-led evaluation', *Canadian Journal of Program Evaluation*, vol. 34, no. 2, pp. 368–384.
- 2019b, *What's working for Māori? A Kaupapa Māori Perspective on the Responsiveness of the Integrated Safety Response Pilot to Māori Synthesis Evaluation Report*, Prepared for Joint Venture Business Unit.
- and McKegg, K. 2018, 'Values and Culture in Evaluative Thinking: Insights from Aotearoa New Zealand', *Evaluative Thinking. New Directions for Evaluation*, no. 158, pp. 93–107.
- Weijer, C., Grimshaw, J.M., Taljaard, M., Binik, A., Boruch, R., Brehaut, J.C., Donner, A.D., Eccles, M.P., Gallo, A., McRae, A.D., Saginur, R. and Zwarenstein, M. 2011, 'Ethical issues posed by cluster randomized trials in health research', *Trials*, vol. 12, pp. 1–11.
- Weiss, C.H. 1993, 'Where politics and evaluation research meet', *Evaluation Practice*, vol. 14, no. 1, pp. 93–106.
- 1998, 'Have We Learned Anything New About the Use of Evaluation?', *American Journal of Evaluation*, vol. 19, no. 1, pp. 21–33.
- Westhorp, G. 2014, *Realist Impact Evaluation: An Introduction*, Overseas Development Institute, London.
- What Works Network 2018, *The What Works Network: Five Years On*.
- White, H., Sabarwal, S. and de Hoop, T. 2014, *Randomized Controlled Trials (RCTs)*, UNICEF Office of Research, Florence, Italy.
- WHO (World Health Organisation) 2013, *WHO Evaluation Practice Handbook*, Geneva.
- Williams, B. 2001, *Building Evaluation Capability*, <http://www.bobwilliams.co.nz/ewExternalFiles/capability3-1.pdf> (accessed 2 June 2020).
- Williams, E. 2019, 'The Challenges of Establishing and Growing an Internal Evaluation Unit', presented at Australian Evaluation Conference, Sydney.
- Williams, M. 2018, 'Ngaa-bi-nya Aboriginal and Torres Strait Islander program evaluation framework', *Evaluation Journal of Australasia*, vol. 18, no. 1, pp. 6–20.
- Wilson, J. 2020, 'History', *Te Ara - the Encyclopedia of New Zealand*.
-

-
- Winangali Ipsos Consortium and IPSOS 2017, *Stronger Communities for Children Evaluation Report*, Canberra.
- Wolgemuth, J.R., Savage, R., Helmer, J., Harper, H., Lea, T., Abrami, P.C., Kirby, A., Chalkiti, K., Morris, P., Carapetis, J. and Loudon, W. 2013, 'ABRACADABRA aids indigenous and non-indigenous early literacy in australia: evidence from a multisite randomized controlled trials', *Computers and Education*, vol. 57, pp. 250–254.
- Wong, P. 2013, *Revised Explanatory Memorandum: Public Governance, Performance and Accountability Act Bill 2013*.
- World Bank 2019, *World Bank Group Evaluation Principles*, International Bank for Reconstruction and Development/The World Bank, Washington.
- Wyatt, K. 2019, *Voice Co-Design Senior Advisory Group*.
- Yap, M. and Yu, E. 2016a, *Community Wellbeing from the Ground Up: a Yawuru Example*, Research Report 3/16, Bankwest Curtin Economics Centre.
- and —— 2016b, 'Operationalising the capability approach: developing culturally relevant indicators of Indigenous wellbeing - an Australian example', *Oxford Development Studies*, vol. 44, no. 3, pp. 315–331.
- Yates, B. and Marra, M. 2017, 'Introduction: Social Return on Investment (SROI)', *Evaluation and Program Planning*, no. 64, pp. 95–97.