



Australian Government
Productivity Commission

Indigenous Evaluation Strategy

Productivity Commission
Background Paper

October 2020



The Productivity 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: 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, 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, Background Paper.*

An appropriate reference for this publication is:

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

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).

Foreword

Policies and programs that affect Aboriginal and Torres Strait Islander people are not working as well as they need to. Part of the problem is that there is little evidence about what is working, what isn't, where and for whom. Evaluation can play an important role filling this gap, but regrettably it is often an afterthought and of poor quality. Crucially, Aboriginal and Torres Strait Islander people are rarely asked about what, or how, to evaluate, or what evaluation results mean.

The Productivity Commission has long advocated for better evaluation to inform decisions about policies and programs. We welcomed the opportunity to develop an Indigenous Evaluation Strategy — with Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges at its centre — to improve the quality and usefulness of evaluations, and strengthen the evidence base for government and Aboriginal and Torres Strait Islander decision makers alike.

At a time when relationships between governments and Aboriginal and Torres Strait Islander people are changing, the Strategy aligns with the aims of the National Agreement on Closing the Gap for governments to change the way they work to improve the lives of Aboriginal and Torres Strait Islander people.

There are no simple formulas. Different contexts call for different approaches to evaluation. Engagement and partnership with Aboriginal and Torres Strait Islander people is fundamental to lifting the quality of evaluations, as is planning early so that the right questions are asked and the right data collected. Central support and ongoing monitoring of agencies' progress implementing the Strategy will be vital to building a culture of evaluation, and support learning and accountability.

Michael Brennan
Chair

Romlie Mokak
Commissioner of the Project

Acknowledgements

The Commission is grateful to the people and organisations who contributed to the development of the Strategy, including those who participated in meetings, roundtables and workshops, and those who made submissions.

We are particularly appreciative of the many Aboriginal and Torres Strait Islander people and organisations who shared their experiences and knowledge with us, in submissions, in person on our visits around the country and by videoconference in the latter stages of the project.

We want to thank staff working in Australian, state and territory government agencies for their frank and valuable insights into the current state of evaluation and the Australian Government agencies who responded to our questionnaire.

Commissioner Mokak would also like to express his appreciation to the staff who worked on the project. The team was led by Rosalie McLachlan and Clare Sibly, and included Peter Daniel, Mark Bryant, Danielle Venn, Marcelo Muñoz, Peter Garrick, Jane Moffett, Daniel McDonald, Claire Prideaux, Pia Andriguetto, Phil Smith, Tyler Reysenbach, Bill Henderson, Karen Carmichael and Ingrid Ottaway.

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

10 April 2019

Contents

Foreword	iii
Acknowledgements	iv
Letter of direction	v
Abbreviations	x
Overview	1
1 What we were asked to do and our approach	4
2 The current state of evaluation	5
3 A new approach — an Indigenous Evaluation Strategy	10
4 Review and future governance arrangements	41
1 About the Strategy	45
1.1 Some background to the Strategy	45
1.2 A changing policy landscape	47
1.3 What is evaluation and why evaluate?	50
1.4 Why a whole-of-government Indigenous evaluation strategy?	52
1.5 How we approached the task	55
1.6 Improving lives is the objective	59
1.7 A roadmap for the rest of the paper	71
2 Evaluation in the APS — history and current practice	73
2.1 The ebbs and flows of evaluation history	73
2.2 The current and changing landscape of evaluation	78
2.3 How do Australian Government agencies currently approach evaluation?	82
2.4 What lessons are there for the Indigenous Evaluation Strategy?	95

3	The state of evaluation for policies affecting Indigenous people	97
3.1	What we heard about the current state of evaluation	98
3.2	What policies and programs have been evaluated?	102
3.3	What does the evidence tell us about what works (or does not work)?	113
4	Evaluation types, approaches and methods	129
4.1	Evaluation types, approaches and methods	130
4.2	How do you decide on an evaluation type, approach and method?	152
5	Lifting evaluation quality: Early planning and engagement	163
5.1	What makes for a high-quality evaluation?	164
5.2	Early planning in practice	175
5.3	Effectively engaging with Aboriginal and Torres Strait Islander people	180
6	What to evaluate	205
6.1	Why set evaluation priorities?	206
6.2	Approaches to setting evaluation priorities	209
6.3	Priority-setting under the Strategy	214
6.4	Initial implementation and a process of review	229
7	Embedding a culture of evaluation and learning	231
7.1	What is a culture of evaluation and learning?	233
7.2	Enhancing the use of evaluations across the APS	236
7.3	Creating incentives to produce high-quality and useful evaluations	258
8	Evaluation capability	281
8.1	What is the current state of evaluation capability in the APS?	283
8.2	The skills required to undertake and commission evaluations	285
8.3	Cultural capability to implement the Strategy	290
8.4	Capability to commission high-quality evaluations	301
8.5	Resourcing evaluation	310
8.6	Who should be responsible for building evaluation capability?	313

9	Data for evaluation	317
9.1	How are data currently used in evaluations?	318
9.2	Issues raised about data for evaluation	321
9.3	Improving data for evaluation	333
10	Monitoring and reviewing the Strategy	357
10.1	Monitoring arrangements for the Strategy	358
10.2	Review arrangements for the Strategy	369
11	Bringing it all together: an Indigenous Evaluation Strategy	373
11.1	The key components of the Strategy	374
11.2	What principles should underpin the Strategy?	379
A	Conduct of the project	393
B	Information request to Australian Government agencies	407
C	A history of evaluation in the APS	435
D	International approaches to evaluation	447
	References	473

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	The Australia and New Zealand School of Government
APS	Australian Public Service
ATSCIC	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 for Social Research and Evaluation (New Zealand)
Cwlth	Commonwealth
DEX	Data Exchange
DFAT	Department of Foreign Affairs and Trade
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
EC	Empowered Communities
ECPF	Enhanced Commonwealth Performance Framework
FMIP	Financial Management Improvement Program
FNIHCC	First Nations and Inuit Home and Community Care

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
IAEP	Impact Assessment and Evaluation Plan
IAS	Indigenous Advancement Strategy
IHS	Indian Health Service (United States)
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
OIPC	Office of Indigenous Policy Coordination
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)
RCT	randomised controlled trial
RFQ	request for quote
RFT	request for tender
RSAS	Remote School Attendance Strategy
SPRC	Social Policy Research Centre
TECs	Tribal Epidemiology Centers
TPK	Te Puni Kōkiri (New Zealand)
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 little about their impacts, 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, 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 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 knowledges, and lived experiences needs to be reflected in what is evaluated, how evaluation is undertaken, and the objectives of policies and programs. A strategy embodying these principles will help achieve the outcomes in the new National Agreement on Closing the Gap.
- The Strategy provides principles-based guidance for Australian Government agencies for selecting, planning, conducting and reporting evaluations. It aims to ‘lift the bar’ on the quality of evaluations and improve their usefulness.
- The Strategy’s proposed governance arrangements (which are essential architecture for an effective Strategy) include an Office of Indigenous Policy Evaluation (OIPE) and an Indigenous Evaluation Council (with all Aboriginal and Torres Strait Islander members). The OIPE and the Council would work in partnership to: monitor and report on agencies’ progress implementing the Strategy; identify evaluation priorities and potential cross-agency/topic evaluations; and provide evaluation leadership and guidance.
- A central clearinghouse for evidence on the effectiveness of policies and programs affecting Aboriginal and Torres Strait Islander people would strengthen accountability for undertaking good evaluations and improve diffusion of knowledge.
- 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 implement policies to improve outcomes for people. But there is always some uncertainty about how policies will play out in practice. 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 other 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 such evidence is central to designing effective policies.

Empowered Communities (and many others) told the Commission that what is needed is 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 push for rigorous evaluation of social and health policy has been described as 'the quiet movement to make government fail less often'.¹

To do that, evaluations need to be high quality and useful, and drawing on Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges is an essential ingredient. This is a fundamental shift from the way evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are typically 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)

¹ Leonhardt (2014).

Evaluation is not just about assessing policies against their objectives, it is also about working out the direction of new policies. And that direction needs to give weight to what Aboriginal and Torres Strait Islander people value. If the outcomes of policies are not what is valued by Aboriginal and Torres Strait Islander people then those policies will likely have limited value and little prospect of success.

There is a symbiosis between the Strategy this paper is proposing and the new National Agreement on Closing the Gap (the National Agreement). The National Agreement makes Aboriginal and Torres Strait Islander organisations partners in decision-making processes for Closing the Gap. This better clarifies what is valued by Aboriginal and Torres Strait Islander people in policy design, implementation and outcomes. Evaluation will in turn test whether the policies and programs that result from this new collaborative approach achieve their goals.

1 What we were asked to do and our approach

The Commission was asked to develop a whole-of-government (Australian Government) evaluation strategy for policies and programs affecting Aboriginal and Torres Strait Islander people. 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
- the approach the Commission will use to review agencies' conduct of evaluations against the strategy.

How we approached the task

The objective of any policy 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 a policy outweigh the costs to the Australian community. The Commission’s approach to developing the Indigenous Evaluation Strategy (the Strategy) was to look 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).

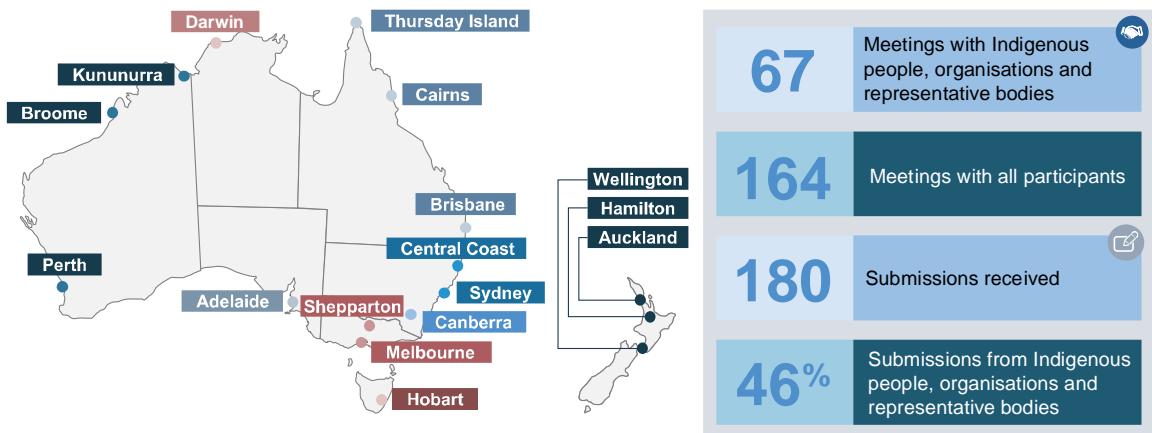
The Strategy is informed by: the literature on good practice evaluation approaches and methods; other evaluation strategies; ethics and research guidelines, standards and strategies;

and governance arrangements for evaluation adopted in other similar countries (New Zealand, Canada, the United States and the United Kingdom).

We engaged 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 and organisations 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’ — and a workshop on ‘Evaluation Priority Setting’ 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



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 no:

- central unit providing leadership on or overseeing evaluation
- Australian Government-wide evaluation strategy or policy to guide evaluation across the APS

-
- 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)
 - requirement for Australian Government agencies to publish evaluations
 - centralised register or repository for evaluations.

Consequently, 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 Health). Other agencies devolve responsibility for prioritising and managing evaluation to the areas responsible for policies and programs. It is common practice for Australian Government agencies to use external evaluators to evaluate policies and programs.

This ad hoc approach diminishes accountability and undermines evidence-based policy, a point made by the recent Independent Review of the APS. The Review recommended establishing a central evaluation function, however, the Australian Government's response has been modest — a small team established in the Department of Finance to develop evaluation guidelines and to help build evaluation expertise in the APS.

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.

Government agencies also told us 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 pointed to examples where evaluation evidence had informed policy decisions. One example is the School Enrolment and Attendance Measure, which provided social work support, as well as the potential to suspend parents’ income support payments, to ensure children enrolled in and regularly attended school. 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.

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

Participants expressed concerns about the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people in a number of areas:

- the lack of evaluation (including that few evaluations of mainstream policies and programs examine impacts on Aboriginal and Torres Strait Islander people)
- the ad hoc way that decisions about what to evaluate are made
- 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
- evaluation is often an afterthought and not integrated into the policy cycle
- 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 (which means the priorities and perspectives of Aboriginal and Torres Strait Islander people and communities are not reflected in evaluations)
- the usefulness of evaluations, including that evaluations:
 - are often not effective in bringing about change
 - focus more on accountability than on finding ways to improve outcomes for Aboriginal and Torres Strait Islander people
 - 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)

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 ... 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 (52 agencies responded, box 2) align with what participants told us about evaluations affecting Aboriginal and Torres Strait Islander people. We received details on 117 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

-
- only about 30 per cent of evaluations of Indigenous-specific policies and programs reported engaging with Aboriginal and Torres Strait Islander people in evaluation decision making
 - just one of 56 evaluations of mainstream policies and programs involved Aboriginal and Torres Strait Islander people in evaluation planning and decision making
 - more than half of the 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
 - two-thirds of the evaluations used a mix of qualitative and quantitative methods, including surveys, interviews, consultation meetings and focus groups. Qualitative methods were the main method used in about one-fifth 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 or other economic analysis
 - about 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.

Fifty-two agencies responded to the questionnaire — 16 departments and 36 agencies (including four Indigenous-specific agencies). Agencies provided information on 307 evaluations conducted between 2016-17 and 2019-20. Just under three-quarters of the evaluations were published. Agencies also mentioned (but did not provide details on) a further 202 unpublished evaluations.

Of the 307 evaluations provided by agencies, 117 mentioned or provided results for Aboriginal and Torres Strait Islander people (56 evaluations were for mainstream policies or programs and 61 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 — eight per cent of agencies said they ‘always’ published evaluation reports and about 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 related to public and community health services, 13 per cent school education, 12 per cent labour and employment and 10 per cent 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)
- 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
- formal ethics assessment was undertaken for a small proportion of evaluations (about one in four mainstream program evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people, and one in three Indigenous-specific program evaluations).

3 A new approach — an Indigenous Evaluation Strategy

The Strategy’s objective

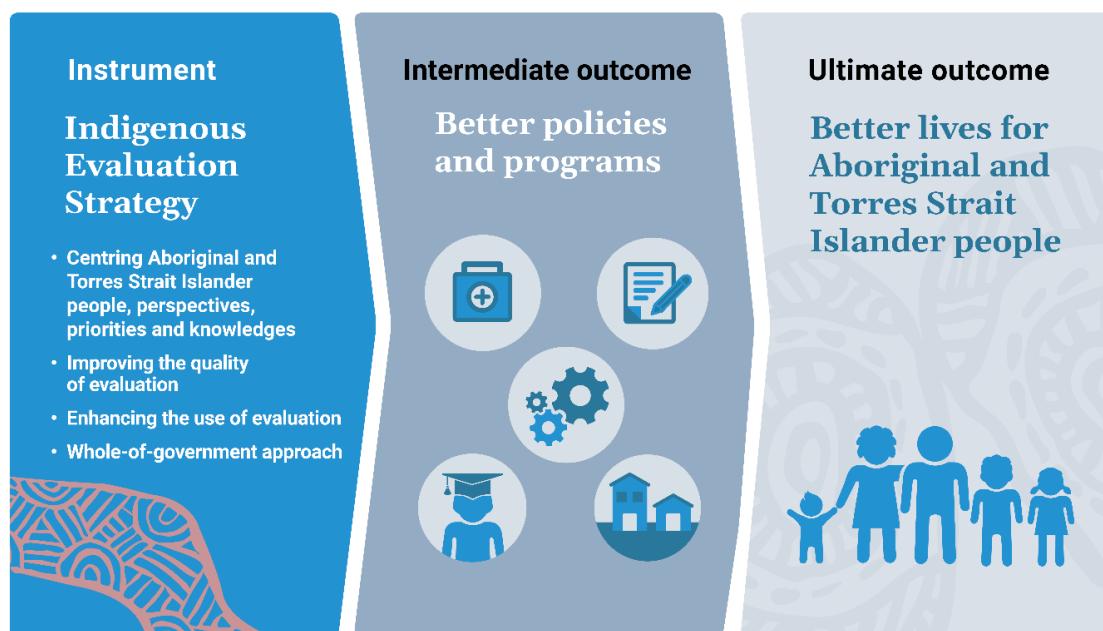
There is widespread agreement that the objective of the Indigenous Evaluation Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people.

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.

There are four ways an Indigenous Evaluation Strategy, via better policies and programs, can 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 all stages of evaluation
- ‘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 an accessible 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 to be partners in the design and conduct of evaluations, 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). Others said the Strategy should reflect the aims of the National Agreement on Closing the Gap — particularly partnerships and shared decision making with Aboriginal and Torres Strait Islander people.

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 and usefulness of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people related to 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 (box 4). This engagement improves evaluations, policy and program design and objectives, and above all, increases the likelihood that policies and programs achieve valued outcomes.

A focus on lived experiences and valued outcomes involves recognising the diverse languages, knowledge systems, 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. For example, kidney dialysis in a hospital (away from the communities of patients) has benefits, but the social costs of patients being separated from Country (including disruption to kinship and cultural ties, and not being able to fulfil cultural commitments) may outweigh these benefits. A mobile dialysis truck that visits remote communities allows patients to have dialysis on Country and achieves both benefits.

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 partnerships 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)

Liz Curran:

Taxpayer funding is important but ought not be used as a pretext to exclude the voices of those directly affected by the services. It will never be spent wisely if, those experiencing how the service is delivered and delivering the service are not heard. It is they who have the sense of lived experience of poor health, justice, and social and wellbeing outcomes and how they can impact on their lives and what at a local level can be improved or is working or in the process of being about to work. (sub. 2, p. 2)

Box 4 Indigenous knowledges and worldviews

Indigenous or traditional knowledges refers to the complex bodies and systems of knowledge, know-how and practices developed and maintained by Indigenous peoples. These knowledge systems are based on values, beliefs, rituals and community laws and practices. They include concepts and methods for land and ecosystem management.

Indigenous knowledges tend to be collectively owned and are often transmitted orally from one generation to another. They can take the form of stories, songs, beliefs, customary laws and artwork, as well as more technical knowledges and skills. The transmission of traditional knowledges is a practical process of continuing and sharing knowledges, and preserving Indigenous cultures, languages and identities.

The value of Indigenous ways of knowing and doing as an input into policy design and evaluation is increasingly being recognised and acted upon. Some examples include using:

- Indigenous knowledges with artificial intelligence to solve complex environmental management problems (CSIRO, sub. DR131, p. 7)
- Indigenous customary fire management to reduce carbon emissions from wildfires in northern Australia.

However, some participants suggested that Indigenous knowledges are not always heard:

- ‘Aboriginal and Torres Strait Islander people can say things and those listening cannot actually ‘hear’ what is being said’ (Maggie Walter, sub. 112, p. 2).
- The Kimberley Aboriginal Law and Cultural Centre and Cultural Development Network said that a key threshold issue for the Indigenous Evaluation Strategy is ‘the recognition that Indigenous Australians have an existing knowledge that is ignored at best, and in most cases, actively worked against’ (sub. DR129, p. 2).

The Institute for Human Security and Social Change also noted the potential challenges for evaluation when there are different values and worldviews.

... the worldviews and values which are encountered in consultation with Aboriginal and Torres Strait Islander people are often different to the values and worldviews of government and other white Australian stakeholders, creating inherent tensions and challenges for evaluation framing and implementation. (sub. DR119, p. 2)

Partnering with Aboriginal and Torres Strait Islander people can strengthen evaluation capability, facilitate a sense of joint ownership and support self-determination. It can also strengthen trust and the cultural capability of non-Indigenous evaluators and those delivering policies or programs. Strengthening the capability of Aboriginal and Torres Strait Islander people to undertake and participate in evaluations is important for improving the quality of evaluations. As Ernst & Young said:

Having First Nations’ peoples evaluating programs with and for First Nations’ peoples and communities is inherent to self-determination and empowering First Nations’ people. The deep insights and intelligent cultural nuancing First Nations’ evaluators can bring would play a fundamental role in improving the quality of policy and programs, as well as signalling the value government places on this critical aspect of evaluation. (sub. DR140, p. 5)

The evidence from evaluations is that having Aboriginal and Torres Strait Islander people and communities integral in the design, development and delivery of policies and programs is an important element of their success. This involves an approach to engagement that is driven by Aboriginal and Torres Strait Islander people, and that addresses power inequalities and supports shared decision making.

The overarching principle of the Strategy

The overarching principle of the Indigenous Evaluation Strategy should be centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (figure 3).² 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.

Centring is also about valuing Aboriginal and Torres Strait Islander knowledges, cultural beliefs and practices, and building capability among Aboriginal and Torres Strait Islander evaluators, organisations and communities. And it is about non-Indigenous evaluators having the necessary knowledge, experience and awareness of their own biases to work in partnership with, and to draw on the knowledges of, Aboriginal and Torres Strait Islander people. Commenting on working with Indigenous people on evaluations, Wehiipeihana noted that:

... deep cultural knowledge is needed for evaluators to work safely and respectfully with Indigenous communities ... [this knowledge helps evaluators to] respectfully navigate entry, agree on ways of working, including methods and approaches, and support Indigenous people to give expression to their views, values and experiences. (2019a, p. 374)

In the context of mainstream policies and programs, centring is about agencies considering the impacts of policies and programs on Aboriginal and Torres Strait Islander people when designing policies and programs and undertaking evaluations. This includes looking at whether outcomes or access to policies or programs are different for Aboriginal and Torres Strait Islander people than for non-Indigenous people. Even when Aboriginal and Torres Strait Islander people are eligible for a policy or program, there may be cultural, social or economic factors that are barriers to accessing it.

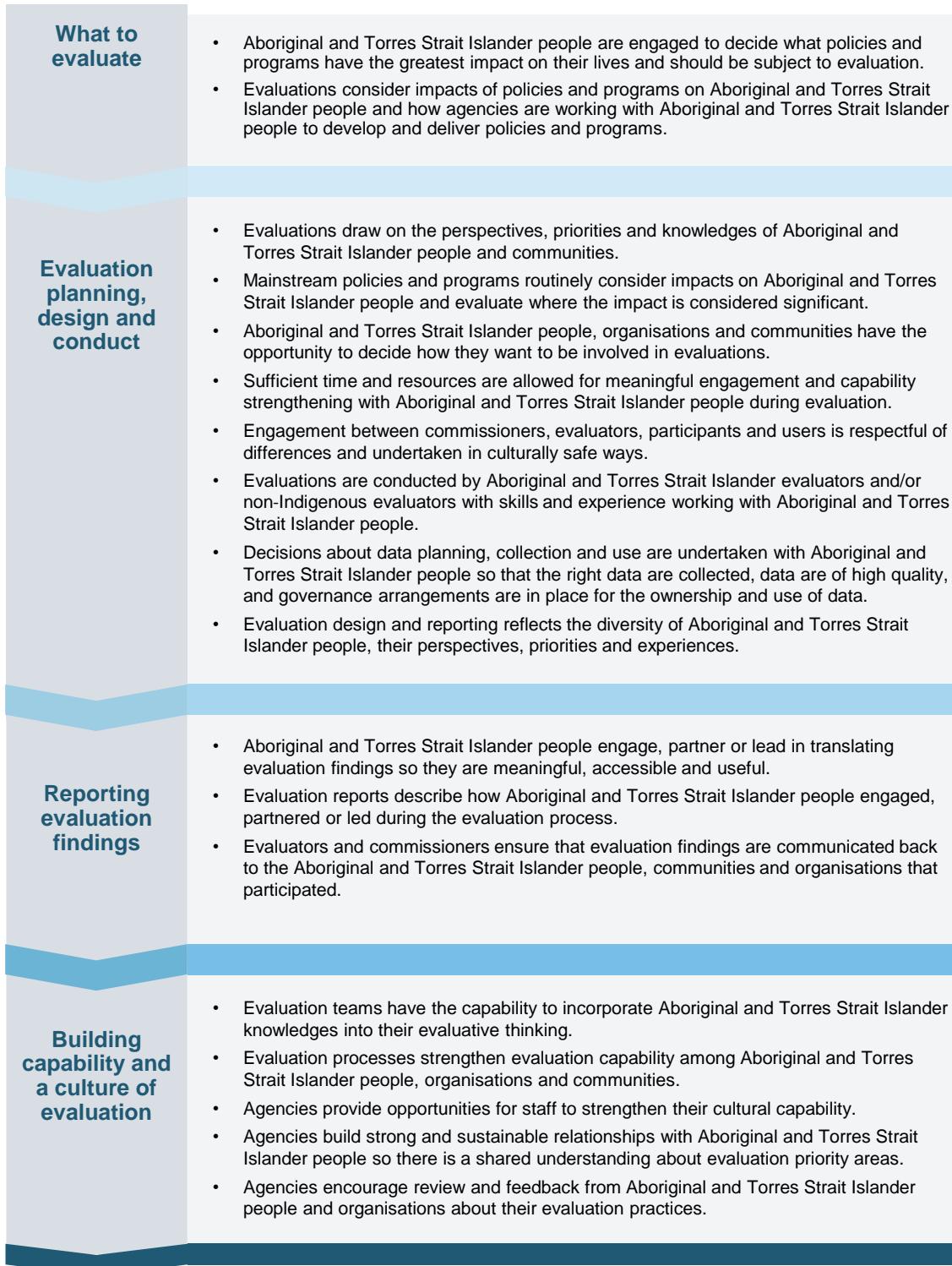
² The overarching principle was supported by Aboriginal and Torres Strait Islander people and organisations as well as government agencies and people doing evaluations.

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



It is important that any potentially different impacts of a policy or program on Aboriginal and Torres Strait Islander people are considered at the policy development/early evaluation planning stage so that: the right data (including baseline data) are collected to track and report on outcomes for Aboriginal and Torres Strait Islander people; program logic reflects on the potential impacts on Aboriginal and Torres Strait Islander people; appropriate evaluation questions are built into the evaluation.

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



The approach to evaluating mainstream policies and programs under the Strategy will depend on the type of policy or program being evaluated, the number or proportion of Aboriginal and Torres Strait Islander people affected by the policy or program, and what is already known about the impacts of the policy or program on Aboriginal and Torres Strait Islander people. Focus groups, case studies or targeted evaluations could be used to examine particular issues affecting Aboriginal and Torres Strait Islander people more closely (a hypothetical example is provided in box 5).

Box 5

Understanding why Aboriginal and Torres Strait Islander people have different outcomes — a hypothetical example

An evaluation of a hypothetical mainstream employment program found that Aboriginal and Torres Strait Islander people had poorer employment outcomes than non-Indigenous people. To understand why this was the case, the agency undertook a further evaluation focusing on the outcomes of the program for Aboriginal and Torres Strait Islander people.

The evaluation used administrative data to look at outcomes for different groups of Aboriginal and Torres Strait Islander people. This analysis found that some of the difference in employment outcomes between Aboriginal and Torres Strait Islander participants and non-Indigenous participants was because of higher caring burdens, and that carers typically had poorer outcomes from the program, regardless of their Indigenous status.

The quantitative analysis also showed that Aboriginal and Torres Strait Islander people had disproportionately poor employment outcomes in some regions. The evaluators conducted case studies of several regions where outcomes were particularly good and poor. Focus groups were held with service providers and program participants in selected regions to better understand regional differences and to collect examples of good practice to share across regions. Surveys were also conducted with participants in the case studies to explore possible unintended consequences (both positive and negative) of the program.

Co-design

Co-design is an important approach to ensuring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are incorporated into evaluations. Co-design reflects higher levels of influence along the engagement continuum (figure 5) — although in practice the term is used more broadly.

In the context of the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, co-design is about partnering with Aboriginal and Torres Strait Islander people and communities to design and conduct an evaluation in a way that is meaningful, respectful, and supports empowerment and ownership.

Figure 5 The continuum of engagement



Co-design usually involves: a design-led process (iterative, person-centred and action oriented); participatory principles (enabling and empowering people affected by a policy or program to make decisions); and tools to get those affected by a policy or program to talk about, enact or propose solutions. An example is provided in box 6.

Box 6 A co-designed evaluation: a trauma healing program

An 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 (box 10) 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 on the factors that cause stress for children 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 experiences and input from students, families and other 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’.

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 will not suit all evaluations. It can be resource- and time-intensive, and those involved need to have access to the information, skills,³ capacities and resources to be able to participate effectively. Government agencies also need the authority, capacity and skills to take an open and flexible approach to evaluation. A reflection from the

³ Including skills in evaluation and data analysis.

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.

Further research and experience is needed to understand how co-design can better work in practice, both in policy and program design and in evaluation itself.

Centring and priority setting

Evaluations should also be undertaken in the areas that are valued most highly by Aboriginal and Torres Strait Islander people. The proposed governance arrangements for the Strategy include an Indigenous Evaluation Council, with all Aboriginal and Torres Strait Islander members, to partner with an Office of Indigenous Policy Evaluation (OIPE) to identify whole-of-Australian Government evaluation priorities (further details are provided below). This is in line with the National Agreement on Closing the Gap, the Australian Government’s policy of working in genuine 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 supported by evidence are more likely to have successful outcomes.

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 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 these 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 programs⁴ 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 also critical for quality evaluations. Building evaluation into the design of policies and programs early helps ensure that the data and evidence needed to answer the evaluation questions and to produce credible results are collected. This includes collecting baseline data, finding an appropriate comparison group and collecting data as policies and programs are implemented.

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 other minority groups) are considered, and the data required to evaluate these effects are collected. While evaluation questions should be guided by early engagement with Aboriginal and Torres Strait Islander people, important questions for evaluations of mainstream policies and programs are:

- Does the policy or program work 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?
- 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?

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, evaluation approaches and methods, and data collection.

⁴ 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

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

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 necessary time and resources are factored into the evaluation for such engagement.

To help embed early planning and engagement for evaluation in the policy cycle (so evaluation is not something that agencies do as an afterthought), the Commission is proposing that an ‘Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plan’ (IAEP) be undertaken for new policy or program proposals developed by Australian Government agencies. The IAEP should include details on:

- the Aboriginal and Torres Strait Islander people, organisations and communities that are expected to be affected by the proposed initiative
- a proposed engagement plan (including timeframes and cultural safety considerations)
- a proposed evaluation approach (including details on the proposed scale of evaluation)
- the data required and how they will be collected
- an estimated evaluation budget.

While IAEPs would not be public documents, published evaluations that are not able to demonstrate how early engagement and a clear program logic informed the evaluation approach would highlight that an agency had not properly considered key aspects of early planning and engagement.

Better quality data collection and use

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.

When we reviewed Australian Government evaluations, we found many evaluations were not able to answer evaluation questions because of limited data (including the absence of baseline data or a comparison group, and/or poor monitoring data that could not be used in the evaluation). Examples include trial data from an Individual Support Placement trial that covered only two years of a three-year trial, and an evaluation of the Cashless Debit Card that was commissioned a few weeks before the program was implemented, which meant there was not enough time for the evaluation to collect adequate baseline data.

Data on Aboriginal and Torres Strait Islander status are also not always collected in administrative datasets and this can mean impacts on Aboriginal and Torres Strait Islander people cannot be evaluated. For example, an evaluation of the National Human Papillomavirus Program was not able to look at 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.

Data planning should be a key focus of early evaluation planning. It is important to ask questions early about what data (including what data capture what matters for Aboriginal and Torres Strait Islander people), sampling strategies and data collection methods are required to measure impacts for Aboriginal and Torres Strait Islander people (an example of adapting data collection processes for local contexts is provided in box 7).

Box 7

Adapting data collection processes for local contexts

Families and Schools Together 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 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.

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.

Considering sampling issues early means that appropriate sampling strategies — such as deliberate oversampling and employing Aboriginal and Torres Strait Islander interviewers as a way of improving response rates — can be chosen and data collected. Oversampling can improve the accuracy of results by increasing statistical power and reducing statistical bias that can be problematic when working with small numbers. It can also counter 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 (box 8).

Engaging Aboriginal and Torres Strait Islander people from the start on data collection, planning and management is also critical for ensuring that evaluations reflect the priorities and knowledges of Aboriginal and Torres Strait Islander people and support Indigenous data sovereignty (that is, the right of Indigenous peoples to exercise ownership over Indigenous data which are about, and may affect, Indigenous peoples both collectively and individually).

Indigenous data sovereignty is put into practice through Indigenous data governance. There are a number of ways evaluations can build in Indigenous data governance, including, for example, Aboriginal and Torres Strait Islander people working on or with an evaluation team to develop data collection procedures and to contextualise and interpret the data that are collected (some examples are provided in box 9).

Box 8 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 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 their 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.

Box 9 Examples of Aboriginal and Torres Strait Islander data governance in evaluations

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.

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.

Indigenous Australians Health Program

This evaluation was designed with a steering group (more than half of the members were Aboriginal and Torres Strait Islander people). In the evaluation itself 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.

Evaluation types, approaches and methods

There are many different evaluation 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 evaluation type and approach used should depend on the:

- questions that need to be answered (table 2)
- kind of policy or program being evaluated
- circumstances under which the policy or program is being implemented
- time and resources allocated to the evaluation
- 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.

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 different approaches (for example, combining experimental or quasi-experimental approaches with theory-based approaches) and mixed methods (including combining quantitative and qualitative methods). Qualitative approaches and data can help answer ‘why’ questions and provide a more complete picture.

The validity of sampling designs and data analysis techniques is 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).

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

Evaluation type	Questions	Purpose
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 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 assess 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.

Are some approaches more or less suited?

Some participants suggested that certain 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.

What makes RCTs different to other evaluation approaches is a randomly assigned control group that allows evaluators to compare the effectiveness of a policy or program against what would have happened had they not changed anything. RCTs are used extensively 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 where policies and programs are designed to deal with multiple and complex issues and where change is expected to occur in the long term.

Many of the less suitable conditions for experimental or quasi-experimental approaches⁵ — such as: it being difficult to disentangle the effects of a single policy or program because there are several other policies in place; a complex relationship between the expected outcome and the policy or program; it not being practical or ethical to create different comparison and target groups; and small sample sizes — will apply to some evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people and communities. RCTs are generally expensive (but they need not be) which means they are usually restricted to high value policies and programs.

Also, because RCTs can hide variations in 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 experimental and quasi-experimental approaches are suitable for assessing the net impact of policies and programs affecting Aboriginal and Torres Strait Islander people (and there are a number of examples).⁶

Certain approaches, such as developmental, realist and participatory (as highlighted by participants) can be particularly suited to evaluating policies and programs as they:

- are better suited to complex problems, adaptive learning and an agile approach to implementing policy
- allow Aboriginal and Torres Strait Islander knowledges, perspectives and world views to be valued and incorporated into the design and delivery of evaluations (including local ways of knowing, being and doing)
- can engage Aboriginal and Torres Strait Islander people throughout the evaluation process, including by using culturally valid methods, such as yarning (storytelling), ganma (knowledge sharing) and dadirri (listening) (box 10)
- can help government agencies build relationships and trust with Aboriginal and Torres Strait Islander people and communities as they work together to find solutions (this aligns with the concept of reciprocity)
- can strengthen the evaluation capabilities of Aboriginal and Torres Strait Islander people and communities (which in turn empowers and facilitates Aboriginal and Torres Strait Islander- and community-led solutions).

⁵ Quasi-experimental methods vary in the extent to which they approach random assignment, with some sophisticated approaches to matching treatment and non-treatment groups, such as propensity score matching.

⁶ For examples, see Goldstein and Hiscox (2018), Ward et al. (2013), Wolgemuth et al. (2013).

Box 10 **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 is known by different names in different Aboriginal cultures, but it translates as 'contemplation'. It requires researchers to continually be reflective of their relationships with others and as an approach is empowering because it allows Aboriginal and Torres Strait Islander voices to be heard.

Developmental evaluation approaches are often applied to innovative contexts as they seek to maximise the success of a policy or program through incremental review and improvement as it is rolled out. These approaches 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 (the 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 11. This evaluation sought to understand how families engaged with, and responded to, the Remote School Attendance Strategy (this initiative 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). The evaluation found that a one-size-fits all approach is unlikely to be effective in improving school attendance in remote communities. However, it also found that when staff were appropriately trained, were 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.

Ethical practice for quality evaluations

Ethics in evaluation is concerned with good conduct. Applying ethical standards can improve the quality and consistency of evaluation and ensure 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 undertaken in an ethical way.

Ethical practice during evaluation should be guided by existing ethical guidelines for research involving Aboriginal and Torres Strait Islander people. Formal review by an ethics committee will be required for some, but not all, evaluations conducted under the Strategy. However, all evaluation projects should include a systematic and well-documented assessment of ethical risks. In cases where an evaluation does not require formal review by an ethics committee, agencies and evaluators should follow ethical practices.

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 critical 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 part of everyday business — 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. Cultural capability, which is a reflective way of doing and thinking (or an attitude towards culture) rather than particular skills or knowledge, is critical for good evaluation processes and valid results.

Box 11 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 Australian and state and territory 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 a one-size-fits-all approach is unlikely to be effective in improving school attendance in remote communities. However, the RSAS was found to be 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.

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 deciding on 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.

Agencies also need to embed cultural capability in their organisations if the Strategy is to be successful — without this, culturally safe evaluations may not be planned for and undertaken. Under the National Agreement on Closing the Gap, Australian governments have committed to transforming government organisations, including by identifying and eliminating racism and embedding cultural safety.

The Strategy's principles-based guidance on how to conduct evaluations could be expected to increase demand for culturally capable evaluation teams⁷. There should also be more 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 (and as pointed out by a number of participants, resources will be needed to draw on such expertise and to support Aboriginal and Torres Strait Islander people partnering with agencies on evaluation). This in turn should strengthen the capability of Aboriginal and Torres Strait Islander people to undertake evaluations.

While responsibility for developing and maintaining evaluation capability should lie with agencies, there should be central support and resources that agencies can access to build or maintain evaluation capability. Central support would also allow for APS-wide learning and the sharing of lessons across agencies.

The Strategy, and the *Guide to Evaluation under the Strategy*, provides guidance on how the principles should be applied in practice, and questions that agencies and evaluators should consider at the various stages of the evaluation process. The Australian Government, as part of its response to the Independent Review of the APS, has also agreed to establish evaluation guidance and an evaluation profession.

In the draft background paper and Strategy, we said the Head of Evaluation Profession should have responsibility for building evaluation capability. However, due to the COVID-19 pandemic, the Government paused its implementation planning of the reform agenda developed in response to the Independent Review of the APS. As such, at least in the interim, we suggest that responsibility for evaluation capability building sit with the proposed Office of Indigenous Policy Evaluation (see below). This would include:

- providing training for commissioning, conducting, and using evaluations of relevant policies and programs
- facilitating an APS-wide community of practice for people who are involved in evaluating policies and programs that affect Aboriginal and Torres Strait Islander people
- establishing processes through which evaluators can seek secondments or other opportunities for mobility to broaden or deepen their experiences

⁷ This includes evaluators, as well as interviewers, surveyors and local researchers.

-
- developing a strategy to build a cohort of Aboriginal and Torres Strait Islander evaluators within the APS (this could include recruitment, 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 or other forms of external oversight). 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, 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 Australian Government has already acknowledged the need for the Strategy to have independent external oversight by giving the Commission the role of monitoring the performance of Australian Government agencies against the Strategy.

The Commission, as a statutory body, has the independence from the day-to-day activity of government necessary to objectively assess agencies’ performance against the Strategy. However, because the Commission does not have powers to direct government agencies in the exercise of their day-to-day functions, the incentive to lift the bar on the quality and usefulness of evaluations will come from the transparency of agencies’ performance.

The Commission was also asked to identify evaluation priorities and to set out its approach to reviewing agencies’ conduct of evaluations against the Strategy. We propose biennial public reporting and these reports will:

- assess the extent to which Australian Government agencies have complied with the Strategy and have implemented the Strategy effectively⁸
- identify good and/or innovative practice in the evaluation of the relevant policies and programs
- assess the extent to which the Strategy has been effective in encouraging high-quality and useful evaluations

⁸ This will include assessing whether agencies have met their responsibilities under the Strategy, covering how they have decided what to evaluate, whether they have prepared and published a rolling Three Year Evaluation Forward Work Plan (action 2 of the Strategy), how agencies have planned, designed and conducted evaluations, and whether they have published evaluation reports (action 7) and provided management responses to evaluations (action 9).

-
- formalise evaluation priorities, including identifying areas for cross-agency evaluations
 - provide recommendations on how the implementation of the Strategy could be improved.

The Strategy has a ‘progression pathway’ for agencies to use to assess their evaluation and engagement practices. It identifies what unsatisfactory, developing, mature and leading practice looks like for each of the Strategy’s principles for the various stages of an evaluation. The pathway will be used to assess agencies’ performance.

Indigenous governance

To align with the Strategy’s overarching principle of centring Aboriginal and Torres Strait Islander people — and the Australian Government’s commitment to work in partnership with Aboriginal and Torres Strait Islander peoples on matters affecting their lives and communities — the oversight arrangements require Indigenous governance to bring an Indigenous lens. An Indigenous lens will bring lived experiences and expertise in the development, delivery and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. The unit within the Commission responsible for overseeing the Strategy will need to partner with Aboriginal and Torres Strait Islander people to determine government-wide evaluation priorities and to monitor the performance of agencies against the Strategy.

Other roles will also be important if the Strategy is to be effective

But if there is only monitoring, without support for agencies to improve their evaluation cultures and practices, there is a real risk that agencies will view the Strategy as little more than a compliance hurdle. Agencies need guidance on how to implement the Strategy. Evaluation leadership will also be important to build a culture that values evidence and innovation, builds evaluative capability, and shares knowledge about what has and has not worked.

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. Evaluation learnings can also be useful for those involved in evaluations beyond the APS, such as state and territory governments, evaluators and Aboriginal and Torres Strait Islander people, organisations and communities.

In the draft background paper and Strategy, we proposed an Office of Indigenous Policy Evaluation⁹ (OIPE), with the following functions:

- have stewardship of the Strategy
- champion the value of evaluation
- advise Australian Government agencies on what the Strategy requires in practice

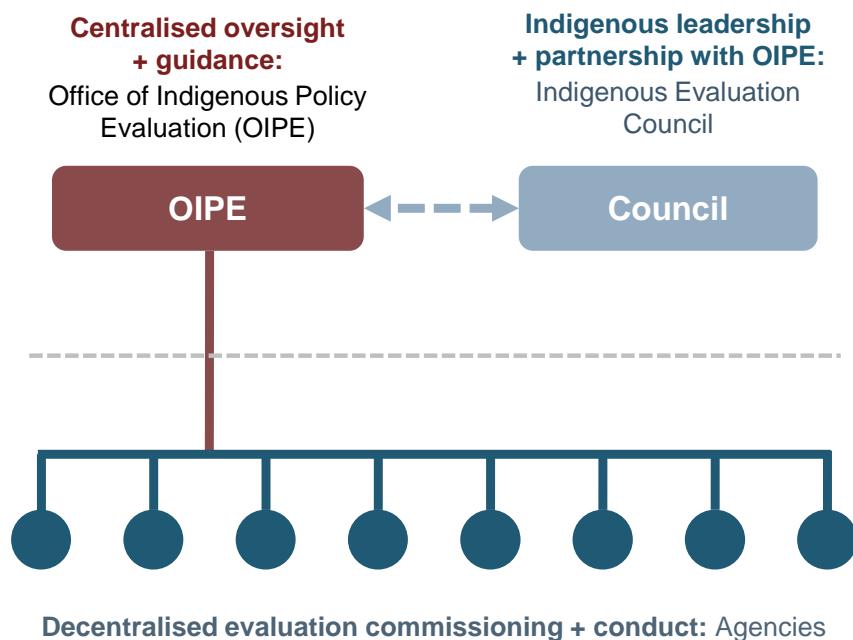
⁹ It was proposed that the OIPE would sit within an existing statutory authority.

- identify evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people and areas for systematic reviews or cross-cutting topics
- oversee agencies' rolling 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.

The OIPE was proposed on the basis that ideally (for reasons of efficiency and effectiveness) all the functions to support the Strategy (and ensure it is effective) would be located in one place. We also sought views at the draft stage on where the OIPE should be located. And in line with the Strategy's principle of centring Aboriginal and Torres Strait Islander people, we proposed an Indigenous Evaluation Council to work with the OIPE.

Locating all the functions in one place means that an Indigenous Evaluation Council (with an all Aboriginal and Torres Strait Islander membership) could partner with and/or advise the OIPE across all the functions (figure 6).

Figure 6 Roles and responsibilities



Some participants supported the OIPE and the Indigenous Evaluation Council (box 12), with some arguing that the National Agreement should be the basis for the formation, roles and operation of the two bodies (with a focus on partnership and shared decision making with Aboriginal and Torres Strait Islander people, and transforming mainstream institutions).

However, others did not support the proposed governance arrangements, arguing that setting up additional government bodies would add unnecessary complexity and administrative burden and would take away resources and responsibility from the community-controlled sector. And others were agnostic about specific governance mechanisms, but stressed that any new arrangements to support the Strategy's implementation must be 'Indigenous-led'.

In terms of where an OIPE might be located, there was some support for the Productivity Commission taking on the additional functions to support the Strategy and housing the OIPE (box 12). Some of the reasons identified for the Commission taking on this role were independence, influence and longevity for the Strategy. Others are that:

- the Strategy's monitoring and public reporting role already sits with the Commission
- there are synergies between the work of the OIPE and the Commission's work on Closing the Gap reporting
- the Commission has an appointed Indigenous Commissioner who could sit on the Indigenous Evaluation Council
- the Commission has well-established, transparent engagement processes in place. It is also a champion for evidence-based policy.

Others suggested the OIPE could sit within the National Indigenous Australians Agency (NIAA), the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), or a non-government Indigenous organisation (box 12). This likely reflects their more established cultural capability and leadership working with Aboriginal and Torres Strait Islander people.

The NIAA currently has as one of its functions analysing and monitoring the effectiveness of policies and programs for Aboriginal and Torres Strait Islander people, including programs and services delivered by other agencies. However, on monitoring agencies' performance against the Strategy, the NIAA may be seen as marking its own homework, and independence is important for credible reporting.

AIATSIS would bring strong Indigenous leadership through its Council, and expertise on ethics and protocols for research with Aboriginal and Torres Strait Islander people to the OIPE. It would also be well placed to make sense of evaluation findings, assess the evidence base and identify priority areas for evaluation. AIATSIS, as an independent statutory authority, could assess agencies' performance against the Strategy; however, it would need to broaden its current research remit (and capacity) to ensure it can oversee evaluation across a wider range of policies and programs.

Box 12 **Comments on the proposed governance arrangements**

Empowered Communities:

If an OIPE is established with functions as proposed it will be crucial that it sits within the broader Productivity Commission. This would both provide a productivity link and better enable the power and influence that it will require to do its job effectively. It will need legislative ‘teeth’ to harness all relevant parties, governments, funded service providers and others, and ensure they follow through on commitments under the Strategy. Establishment in another, even relatively arm’s length, government agency, without its own legislative base and direct productivity link would lessen its strength to act as it will need to, to support Indigenous people and government to close the gap. (sub. DR171, p. 5)

Independent Members of the Indigenous Evaluation Committee, NIAA:

Finding a position for the [Council for Indigenous Evaluation] that is both independent and influential is critically important, but getting this balance right also has implications for the longevity of the IES [Indigenous Evaluation Strategy] and its infrastructure. These governance arrangements, and the IES principles and behaviours to be implemented, need to be sustained over the longer term. We therefore support the proposal that the [OIPE] and the Council be located in an independent but also influential agency such as the Productivity Commission. (sub. DR115, pp. 2–3)

Beyond Blue:

The inclusion of an Indigenous-led evaluation champion to oversee the Strategy and Indigenous-led Evaluation Council to govern the Strategy is a major strength. (sub. DR121, p. 1)

Yulang Indigenous Evaluation:

We strongly support the establishment of an Office of Indigenous Policy Evaluation. ... [and] that the OIPE be situated in an independent agency such as the Productivity Commission. (sub. DR150, p. 16)

National Indigenous Australians Agency:

The success of the Indigenous Evaluation Strategy will be dependent on the establishment of the Centre for Evaluation Excellence *and* the Office of Indigenous Policy Evaluation. Change is needed across the Australian Public Service to both evaluation capability and evaluation use to achieve the intended outcomes of the Indigenous Evaluation Strategy. Particularly important to the success of both bodies will be the appointment of an ‘eminent’ person to lead each one. This will increase the likelihood of success in influencing change across the Australian Public Service. (sub. DR161, p. 6)

Lowitja Institute:

The recommendation to create the Office for Indigenous Policy Evaluation (OIPE) is likely to be controversial because it potentially creates another government body with little power and authority over other agencies to enforce the principles and practices outlined in the Strategy. Our view is that an Aboriginal and Torres Strait Islander governed body to drive the Strategy is important ... we support establishing strong Aboriginal and Torres Strait Islander governance and advisory mechanisms to support the delivery of the Strategy. (sub. DR157, p. 4)

National Health Leadership Forum:

The NHLF supports the proposal for an overseeing role of the IES [Indigenous Evaluation Strategy] across Commonwealth agencies, however it is not fixed on the [OIPE/Council] structure, but rather whatever the structural form, it must be Indigenous led. (sub. DR146, p. 3)

The National Centre of Indigenous Excellence supported AIATSIS housing the evaluation clearinghouse and said:

The clearinghouse must always stay Indigenous. In order to best position the OIPE to help drive organisational change, this function should be incorporated within the OIPE, or at least responsibility for the coordination of this activity should be vested with them, with appropriate funding being provided to support the delivery of the activity. (sub. DR158, p. 22)

A non-government Aboriginal and Torres Strait Islander organisation would also bring valuable insights on what centring means for evaluating policies and programs, and connections with Aboriginal and Torres Strait Islander people, organisations and communities. It would also be well placed to make sense of evaluation findings and identify priority areas for evaluation. However, a non-government Indigenous organisation would not have the independence of a statutory authority to assess agencies' performance (or the stability of a statutory authority). A non-government organisation is also likely to have to build expertise in performance monitoring and evaluation of government agencies.

A second-best option is to separate out some of the functions of the OIPE. For example, the evaluation unit within the Department of Finance could undertake the role of evaluation champion and provide advice to agencies about applying the Strategy in practice. The NIAA could also take on the capability building role in collaboration with the Australian Public Service Commission.

However, as noted earlier, the first best option is to have all the proposed functions to support the Strategy in one place. An OIPE and an Indigenous Evaluation Council working together are considered essential architecture for an effective Strategy.

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 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 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 (see below), 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 currently, 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, a summary report should 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 relevant policies and programs.

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 (the clearinghouse could sit within the OIPE or another agency, such as the Australian Institute of Health and Welfare or AIATSIS). 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, the accumulation of consistent evidence over time and across different contexts gives policy makers greater confidence in evaluation results.

Promoting a whole-of-government approach to priority setting

To help guide agencies' efforts on what to evaluate under the Strategy, the Commission proposes that the OIPE, working in partnership with the Indigenous Evaluation Council, should establish and maintain a formal set of government-wide evaluation priorities.

In the meantime, the priority reform areas and socio-economic outcomes identified in the National Agreement on Closing the Gap provide a sound basis for an interim set of priority evaluations (figure 7). Aligning evaluation priorities with the National Agreement was endorsed by many participants.

Under the proposed interim government-wide evaluation priorities, the four priority reforms under the National Agreement will help identify what actions within (and across) the identified priority policy domains should be prioritised for evaluation.

Figure 7 Proposed interim government-wide evaluation priorities

Based on the National Agreement on Closing the Gap

The Strategy should prioritise evaluation of government efforts ...			
... delivered as part of a formal partnership and/or shared decision making arrangement (Priority 1) that build up the Aboriginal and Torres Strait Islander community controlled sector (Priority 2) (particularly mainstream ones) that eliminate racism, embed and practice meaningful cultural safety, deliver services in partnership with Aboriginal and Torres Strait Islander people (Priority 3) involved in data sharing arrangements (Priority 4) ...
... and focused across (and within) the following policy domains:			
<ul style="list-style-type: none">• Education• Housing• Health and wellbeing	<ul style="list-style-type: none">• Land and water• Justice• Culture and languages	<ul style="list-style-type: none">• Safety (Families, children and youth)• Employment (Economic development)	

It is expected that government-wide evaluation priorities — including opportunities for cross-agency evaluations — will continue to be better defined over time with better information on:

- the implementation of policy and reform priorities included in the National Agreement on Closing the Gap
- the evaluation and policy landscape of agencies, as they report more and better quality information through rolling Three Year Evaluation Forward Work Plans and completed evaluation reports on key policies and programs
- other evaluation priorities identified during the monitoring process.

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 priorities, and assess whether they should be prioritised for evaluation.

Under the Strategy, departments would 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 would 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.

4 **Review and future governance arrangements**

Reviewing the Strategy

While the OIPE would monitor the performance of the Strategy, it is also important to have a separate process for reviewing the Strategy. We propose that the Australian Government commission an independent review of the Strategy five years after it is endorsed. The review should be led by an Aboriginal and Torres Strait Islander person and assess:

- whether the Strategy’s principles remain fit-for-purpose
- the extent to which the Strategy has been effective in encouraging higher quality and more useful evaluations (this includes assessing the effectiveness of the Strategy’s actions)
- the extent to which oversight arrangements remain fit-for-purpose, including by assessing the performance of those overseeing the Strategy and whether changes to the oversight arrangements are required.

A Centre for Evaluation Excellence

The case for central evaluation leadership and external oversight is wider than just evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. As such, over the longer term (five years), a new independent Australian Government agency — a Centre for Evaluation Excellence (CEE) — should be established (figure 8).

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: <ul style="list-style-type: none"> • Guidance material and advice to agencies on good practice evaluation governance and standards • Facilitate evaluation community of practice across APS; coordinate opportunities for upskilling in evaluative capability 	<ul style="list-style-type: none"> • Guidance and advice on Indigenous Evaluation Strategy governance and standards, and good practice • Cultural capability (so evaluations centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges) • Develop strategy for building cohort of Aboriginal and Torres Strait Islander evaluators
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 Evaluation 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

The CEE would have similar leadership and oversight roles as the OIPE, however, it would lead efforts to improve evaluations of social and health policies across the APS as a whole. The Commission also proposes that when the CEE is established, the OIPE could 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 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).

1 About the Strategy

The Australian Government provides a wide range of policies and programs that affect 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 incorporating 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)

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 2010, 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 at the Productivity Commission's 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, Aboriginal and Torres Strait Islander communities (PC 2013). And one of the key messages of the Commission's *National Indigenous Reform Agreement Performance Assessment 2013-14* was that while there was extensive reporting on outcomes for Aboriginal and Torres Strait Islander people, more attention and resources were needed for policy evaluation (PC 2015).

In the 2017-18 Federal 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, programs and communities. 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 2017).

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 — schooling, employment 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 (2020b).

More about what the Commission was asked to do

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

1.2 A changing policy landscape

In July 2020, the Joint Council on Closing the Gap signed a new National Agreement on Closing the Gap (the National Agreement) (2020). The National Agreement provides an overarching framework for developing and implementing policies and programs that impact on Aboriginal and Torres Strait Islander people. The National Agreement builds on, and replaces, the National Indigenous Reform Agreement. It is significant in that for the first time representatives of Aboriginal and Torres Strait Islander people (the Coalition of Aboriginal and Torres Strait Islander Peak Organisations, which is made up of about 50 Aboriginal and Torres Strait Islander community controlled peak organisations and members (Coalition of Peaks 2020b)) are parties to the National Agreement.

The National Agreement outlines four priority reform areas for joint national action, and establishes 16 socio-economic targets¹⁰ that focus on improving the outcomes experienced by Aboriginal and Torres Strait Islander people (box 1.2).

¹⁰ Additional socio-economic targets relating to family violence, access to information, community infrastructure and inland waters are being developed.

Box 1.2 The National Agreement on Closing the Gap

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 initial Closing the Gap targets expired without being met. Just two of the continuing targets, early childhood education and Year 12 attainment, remain on track.

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 Partnership Agreement sets out how governments and the Coalition of Peaks will work together over the next ten years on Closing the Gap.

In July 2020, a new National Agreement on Closing the Gap (the National Agreement) was signed by Australian governments and the Coalition of Peaks. The National Agreement means that, for the first time, Aboriginal and Torres Strait Islander representatives have been full partners in the (previously government-only) decision-making process for Closing the Gap. In accordance with this new approach, the National Agreement now addresses both change needed in government processes as well as indicators of wellbeing for Aboriginal and Torres Strait Islander people.

Under the National Agreement there are four ‘Priority Reform Areas’. These Priority Reform Areas ‘change the way governments work to accelerate improvements in the lives of Aboriginal and Torres Strait Islander people’ (p. 5). The priority reform areas are:

1. **Formal partnerships and shared decision making** which involves parties committing ‘to building and strengthening structures that empower Aboriginal and Torres Strait Islander people to share decision-making authority with governments to accelerate policy and place-based progress against Closing the Gap’ (p. 5). As part of this, governments have committed to establishing policy and place-based partnerships which respond to local priorities, and to review new and existing partnership arrangements to assess how they meet the elements of strong partnerships, as set out in the National Agreement.
2. **Building the community-controlled sector** to deliver services to support Closing the Gap, which includes having governments implement measures to increase the proportion of services delivered by Aboriginal and Torres Strait Islander organisations and report on actions taken to strengthen the community-controlled sector.
3. **Transforming government organisations** to improve accountability and respond to the needs of Aboriginal and Torres Strait Islander people. As part of this, governments have committed to identify and eliminate racism; embed and practice cultural safety; deliver services in partnership with Aboriginal and Torres Strait Islander organisations; communities and people; increase accountability through transparent funding allocations; support Aboriginal and Torres Strait Islander cultures; and improve engagement with Aboriginal and Torres Strait Islander people.
4. **Shared access to data and information at a regional level** to support Aboriginal and Torres Strait Islander communities and organisations to support the achievement the first three priority reforms.

The National Agreement also sets out 16 socio-economic targets to be pursued. They are to:

1. Close the Gap in life expectancy within a generation, by 2031.
2. By 2031, increase the proportion of Aboriginal and Torres Strait Islander babies with a healthy birthweight to 91 per cent.

(continued next page)

Box 1.2 (continued)

3. By 2025, increase the proportion of Aboriginal and Torres Strait Islander children enrolled in Year Before Fulltime Schooling (YBFS) early childhood education to 95 per cent.
4. By 2031, increase the proportion of Aboriginal and Torres Strait Islander children assessed as developmentally on track in all five domains of the Australian Early Development Census (AEDC) to 55 per cent.
5. By 2031, increase the proportion of Aboriginal and Torres Strait Islander people (age 20–24) attaining year 12 or equivalent qualification to 96 per cent.
6. By 2031, increase the proportion of Aboriginal and Torres Strait Islander people aged 25–34 years who have completed a tertiary qualification (Certificate III and above) to 70 per cent.
7. By 2031, increase the proportion of Aboriginal and Torres Strait Islander youth (15–24 years) who are in employment, education or training to 67 per cent.
8. By 2031, increase the proportion of Aboriginal and Torres Strait Islander people aged 25–64 who are employed to 62 per cent.
9. By 2031, increase the proportion of Aboriginal and Torres Strait Islander people living in appropriately sized (not overcrowded) housing to 88 per cent.
10. By 2031, reduce the rate of Aboriginal and Torres Strait Islander adults held in incarceration by at least 15 per cent.
11. By 2031, reduce the rate of Aboriginal and Torres Strait Islander young people (10–17 years) in detention by at least 30 per cent.
12. By 2031, reduce the rate of over-representation of Aboriginal and Torres Strait Islander children in out-of-home care by 45 per cent.
13. A significant and sustained reduction in violence and abuse against Aboriginal and Torres Strait Islander women and children towards zero.
14. Significant and sustained reduction in suicide of Aboriginal and Torres Strait Islander people towards zero.
15. By 2030, a 15 per cent increase in Australia's landmass subject to Aboriginal and Torres Strait Islander people's legal rights or interests. By 2030, a 15 per cent increase in areas covered by Aboriginal and Torres Strait Islander people's legal rights or interests in the sea.
16. By 2031, there is a sustained increase in number and strength of Aboriginal and Torres Strait Islander languages being spoken.

Each party to the National Agreement will develop an implementation plan. The Joint Council will monitor implementation and may provide advice on implementation and how parties can best work together to achieve shared outcomes. Parties will also report annually on their progress on the National Agreement and their implementation plans.

The Productivity Commission will maintain a dashboard for data and associated supporting material to inform the reporting on progress. The Commission will also review progress every three years, which will include analysis of progress against the priority reforms, targets, indicators and trajectories, and examine the factors contributing to progress. Each Commission review will be followed by an independent Aboriginal and Torres Strait Islander-led review.

Sources: Australian Government (2020b); Joint Council on Closing the Gap (2019, 2020).

Because the National Agreement will provide the basis for new policies and programs affecting Aboriginal and Torres Strait Islander people, evaluation that is high-quality, useful and that incorporates the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people will be important for informing how to achieve the commitments under the National Agreement.

Recognising the importance of the National Agreement for shaping policies affecting Aboriginal and Torres Strait Islander people going forward, the Strategy will be a useful tool for the Australian Government (and also other governments and other parties) to draw on and complement the other actions that they will take to implement the National Agreement.

The 2019 Independent Review of the Australian Public Service (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).

And the Australian Government recently committed \$7.3 million for a co-design process to develop options for a ‘voice to government’ for Aboriginal and Torres Strait Islander people 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)

1.3 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 these policies and programs 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 (box 1.3). Reasons for undertaking evaluation 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.

Box 1.3 **Evaluation and research — what is the difference?**

Both evaluation and research are important tools for evidence-based policy. Both uncover learnings in a systematic way, where the area of inquiry is bound to a question, hypothesis or knowledge gap. However, evaluation and research often (but not always) differ by what is examined, who decides what to examine, and impact.

Evaluation identifies value from actions (policies and programs), and as such, is seen as more interested in specific, applied knowledge. It is undertaken to test the effectiveness of a specific program or model, with its questions typically originating with stakeholders and intended users of evaluation findings. The value from evaluation comes from its usefulness to improve policy or program effectiveness.

Research, on the other hand, is often driven by curiosity around a particular problem or topic – with the nature of the research informed by the priorities of individuals (such as individual researchers), communities (which is often important in research involving Aboriginal and Torres Strait Islander communities, where research is often viewed as being collectively owned), or through combinations of both. Research is typically undertaken to test theories and produce generalisable findings.

Evaluation often seeks to quantify the effects of an intervention based on research knowledge. For example, research may show that vaccinated children do not get a particular disease, whereas an evaluation may seek to quantify the benefits of a vaccination program (in terms of improved life outcomes and cases avoided). An evaluation does not go back and re-prove what research has found.

Sources: adapted from Patton (2017); Rogers (2014).

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 if findings are used to support 'learning by doing' and stakeholders work together to find ways to achieve agreed outcomes.

Many participants 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 1.4). Others pointed to some of the consequences of not evaluating or of undertaking poor quality or poorly focused evaluations, including ad hoc policy decision making, not knowing whether the policy or program 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 1.4).

1.4 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)

In a report on *Mapping the Indigenous Program and Funding Maze*, the Centre for Independent Studies 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)

Box 1.4 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)

Ernst & Young:

Evaluators and agencies responsible for delivering evaluations have a unique and substantial role to play in improving the life-outcomes for First Nations' peoples and communities by informing policy, program and service delivery decisions with the highest quality of insight and assessment. (sub. DR140, p. 2)

Independent Members of the National Indigenous Australians Agency Indigenous Evaluation Committee (NIAA IEC):

The NIAA IEC strongly supports the ambition of the Draft Indigenous Evaluation Strategy (IES). We echo the IES argument that the within-Department embedding of rigorous and appropriately developed evaluation processes is essential if Indigenous policy and programs in Australia are to deliver better outcomes. (sub. DR115, p. 1)

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 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], as well as an over-reliance on anecdotal evidence' (Hudson 2017, p. 13).

The 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 systems. The 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, as discussed earlier, and in keeping with a commitment by the Australian Government to work differently with Aboriginal and Torres Strait Islander people, 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 many Aboriginal and Torres Strait Islander people's views about evaluation. The Strategy should add further impetus to these changing views.

The Strategy is for everyone involved in the evaluation of Australian Government policies and programs affecting Aboriginal and Torres Strait Islander people. The Strategy could also have implications for state, territory and local governments, given the significant crossovers (including how policies and programs are administered and funded) between Australian, state, territory and local governments for policies and programs that affect Aboriginal and Torres Strait Islander people, and the aims of the National Agreement.

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. This includes 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 1.5 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)

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 ... 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)

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)

1.5 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 APS: 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? What feedback loops exist within and across government for using evaluation findings in policy and program design? How are the results of evaluations and subsequent actions communicated to evaluation users and Aboriginal and Torres Strait Islander people? Is there a role for greater oversight of evaluations?
- what data, capabilities and resources are required to produce and use high-quality evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people?

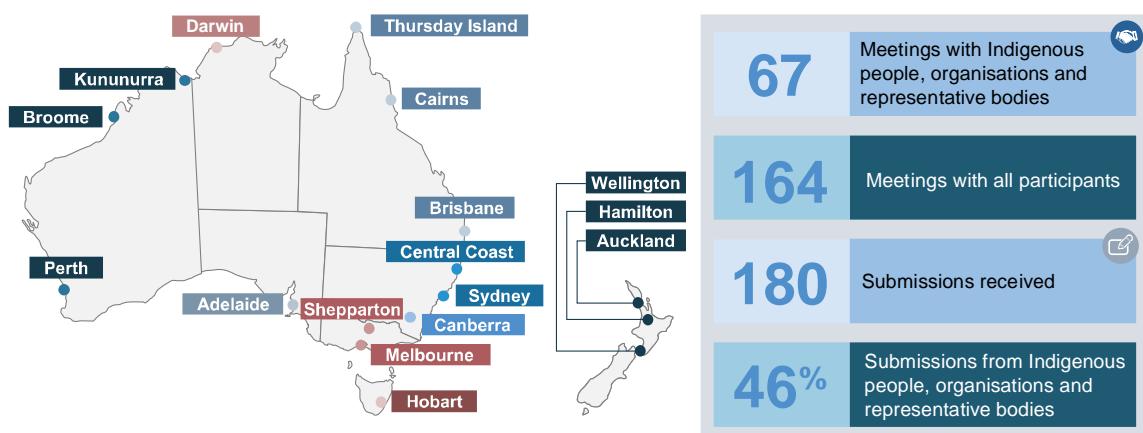
Information from a range of sources was used to help answer these questions.

We received 180 submissions (including four oral submissions) and four brief comments (figure 1.1). We sent a questionnaire to all Australian Government agencies asking for information on their evaluation policies and practices and for any evaluations they had undertaken in the previous three years (for more information on the questionnaire see chapter 3 and appendix B).

The Strategy was also informed by the literature on good practice evaluation approaches and methods, other evaluation strategies, 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 participants¹¹ suggested that Aboriginal and Torres Strait Islander people should be co-designing the Strategy. There were, however, some constraints around co-designing the Strategy as the Australian Government asked the Commission to develop the Strategy. That said, we engaged widely when developing the Strategy, including with many Aboriginal and Torres Strait Islander people and organisations (figure 1.1). The Commission was also cognisant of the changing policy landscape, including the formation of the Joint Council on Closing the Gap and the development of the National Agreement. These processes informed and influenced the development of the Strategy.

Figure 1.1 The Commission engaged widely



The Commission also engaged with Australian, 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 Strategy. We held two roundtables — one on ‘Objectives, principles and defining success’ and another on ‘Culture, capability and governance’, as well as a ‘Priority Setting Workshop’ — to test our thinking on aspects of the Strategy with Aboriginal and Torres Strait Islander organisations, academics and government agencies. Appendix A has a list of project participants and submissions.

The Strategy will be a ‘living’ document that is subject to ongoing monitoring and 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 monitoring and review process.

¹¹ For example: Aboriginal Affairs NSW (sub. 70, p. 2); Generation One, Minderoo Foundation (sub. 5, p. 1); and National Aboriginal Community Controlled Health Organisation (sub. 95, p. 2).

What policies and programs are covered by the Strategy?

The Strategy covers all Australian government policies and programs that affect Aboriginal and Torres Strait Islander people — that is, both ‘mainstream’ policies and programs (universal policies and programs for all Australians) and Indigenous-specific programs (policies and programs specifically for Aboriginal and Torres Strait Islander people). 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 policies and 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 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 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 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

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)

Queensland Family and Child Commission:

Aboriginal and Torres Strait Islander peoples are affected by mainstream policies and programs as much as any other person in Australia. Adverse impacts of whole of population policy and programs that do not understand and respond to the rights, needs and circumstances of First Nations People are demonstrated in the disproportionate rates of involvement in statutory systems and in the disparity that persists across a range of social determinants far beyond the scope of the Closing the Gap target areas. While there is a compelling case for change in the design, delivery and evaluation of targeted supports and services, the impact of all government policies and programs should be subject to evaluation and scrutinised for impacts on outcomes for Aboriginal and Torres Strait Islander peoples. (sub. DR136, p. 6)

1.6 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.

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 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).

Some participants said historical events (box 1.10)¹² (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)

¹² A more detailed discussion on the history of Aboriginal and Torres Strait Islander people since European settlement can be found in chapter 1 of the *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 Aboriginal and Torres Strait Islander 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 Aboriginal and Torres Strait Islander 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 affecting 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 2005, 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 Aboriginal and Torres Strait Islander people

A person's view of wellbeing can be influenced by their values, beliefs and social norms (Wehipeihana 2019a, p. 376). As such, 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)

Ernst & Young also commented that:

... the concept of wellbeing needs to be carefully considered in the context of First Nations' evaluation. We know that core to First Nations' culture is the concept of 'community wellbeing', with wellbeing of the individual being intrinsically linked to collective wellbeing. It is our view that in order to effectively define or make an assessment on wellbeing, it must be clearly defined and understood in the context of the First Nations' communities relevant to the evaluation. (sub. DR140, p. 3)

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 was launched by National Aboriginal and Torres Strait Islander Leadership in Mental Health 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 Aboriginal and Torres Strait Islander 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 Aboriginal and Torres Strait Islander wellbeing (ABS 2010).

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).

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 ways 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)

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



Source: ABS (2010).

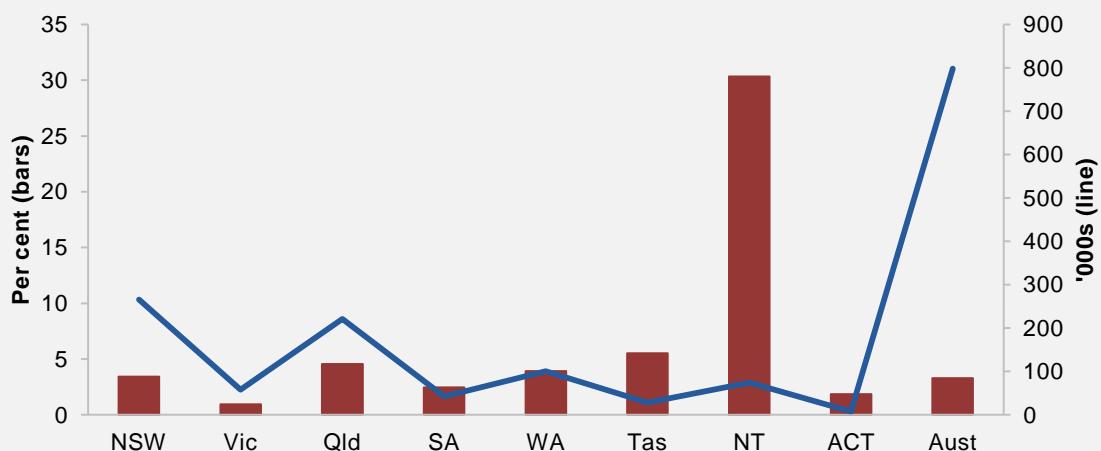
The Aboriginal and Torres Strait Islander population is also geographically diverse (box 1.12). While the conversation about evaluating policies and programs affecting Aboriginal and Torres Strait Islander people is often focused on those living in remote communities, most of the Aboriginal and Torres 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

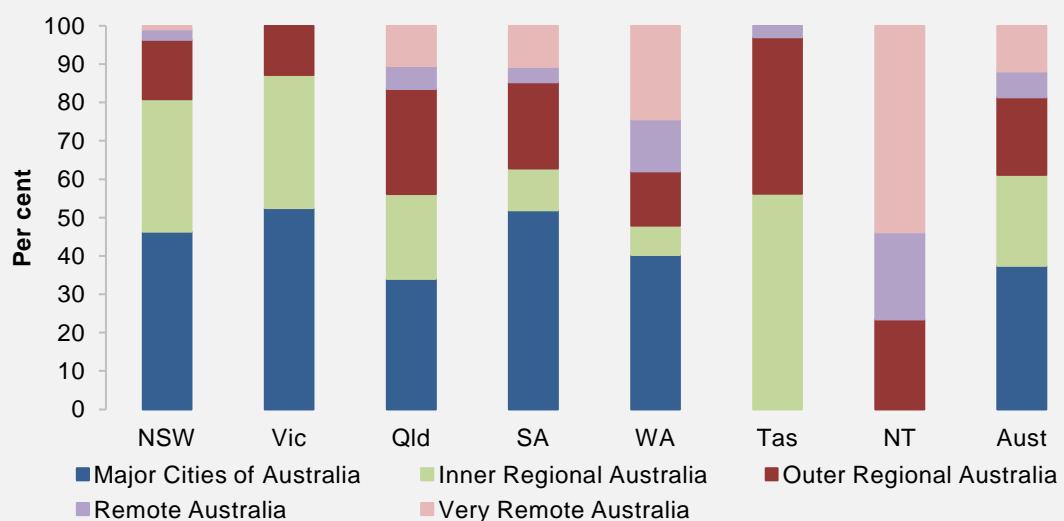
Aboriginal and Torres Strait Islander people make up about 3.3 per cent of the total Australian population. The proportion of Aboriginal and Torres Strait Islander people in each state and territory varies from less than 1 per cent of the population in the Australian Capital Territory and Victoria to 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 per cent). While 18 per cent of the Aboriginal and Torres Strait Islander population live in remote areas, this proportion is 77 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 were not available for the Australian Capital Territory).



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).

The Strategy needs to acknowledge the diversity of Aboriginal and Torres Strait Islander people, 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, 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 embedding Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in policy and program evaluation planning, design, implementation and priority setting (box 1.13).

Many participants called for the Strategy to recognise 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. The declaration outlines minimum standards for the ‘survival, dignity and wellbeing of the indigenous people of the world’ and the right to self-determination.

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)

National Centre of Indigenous Excellence:

The role of culture in evaluation is crucial as it underpins values, processes, findings and, ultimately, outcomes. Within evaluation, the processes of information exchange, interpretation and application of knowledge are significantly influenced by the cultures of evaluation participants, including the evaluator. Within this Strategy, there must be a shift to valuing and centring Indigenous knowledges and perspectives in evaluation design and implementation. It must reflect the ways of knowing, doing and being of the Communities that the program or policy is designed to serve. It is impossible for an evaluation to be meaningful to a Community if the worldviews that underpin and impact on the evaluation are not expressly acknowledged and questioned. (sub. DR158, pp. 7–8)

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

For decades, Aboriginal and Torres Strait Islander people have 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).

While self-determination means different things to different people, the National Health and Medical Research Council explained the right to self-determination to mean that Aboriginal and Torres Strait Islander people 'have the freedom to live well and to live according to their values and beliefs' (NHMRC 2020). There are four main ways 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 Aboriginal and Torres Strait Islander people's priorities, perspectives and knowledges are central to evaluations.

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).

A common theme in the 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.

1.7 A roadmap for the rest of the paper

The next chapter looks at the evaluation environment across the Australian Public Service, and chapter 3 discusses the current state of evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.

Chapter 4 looks at the various types, approaches and methods of evaluation. Chapter 5 discusses what makes for high quality evaluation, with a focus on early planning and engagement.

Evaluation priorities, and how they should be determined, are examined in chapter 6. Chapter 7 looks at governance arrangements to support a culture of learning and evaluation within agencies, and chapter 8 at how evaluation capability might be strengthened. Chapter 9 looks at data for evaluation.

Arrangements for monitoring and reviewing the Strategy are discussed in chapter 10 and chapter 11 outlines the key components of 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 is currently no formal requirement 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 reports or 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 (with some pockets of good evaluation practice).

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 implement 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 flows 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 major reports in the 1970s — the Coombs¹³ 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’ (RCAGA 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)

In 1979, a report by the Senate Standing Committee on Social Welfare, *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 policies and programs 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 from 1979) to a new, quasi-independent body, the Office of Evaluation and Audit (box 2.1).

¹³ 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).

Its legislated remit included evaluating and auditing the operations of the Aboriginal and Torres Strait Islander Commission (ATSIC) regularly (ss. 75-76), with the Act requesting that the director of the OEA — in overseeing the operations of the Office — have regard to the desirability of evaluating and auditing programs at least once every three years (s. 78).

The OEA was independent from the ATSIC Board of Commissioners and reported directly to the Parliament.

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

From its inception, evaluation activity within the OEA 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 were not integrated into agencies' financial decision making, 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 the first whole-of-government evaluation strategy in 1988.

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 — and would later provide a biannual register of published evaluations (box 2.2). Evaluation units were set up in most departments.

Box 2.2 Registers and clearinghouses add value to evaluation activity

To improve access to evaluation and research findings, governments often use registers, repositories and clearinghouses. Contemporary examples include the Research and Evaluation Register maintained by the Victorian Department of Education and Training, and the Indigenous Justice Clearinghouse maintained by the Australian Institute of Criminology and the Council of Attorneys-General.

In the early 1990s, agencies were encouraged to publish evaluation findings in their annual reports. However, this was considered inadequate to inform public sector managers and evaluators about the information that was becoming progressively available, so in 1992, the Department of Finance produced a register of published evaluations. The register was updated twice a year and communicated evaluation findings to agencies and the public while also providing a means to disseminate evaluation strategies and methodologies between agencies. The register appears to have been discontinued in 1996 — alongside other Department of Finance evaluation oversight activities — and no such register of evaluation findings has existed since.

A clearinghouse can act as both a repository of existing research, and as a source of further analysis and translation of that research. The Closing the Gap Clearinghouse — jointly funded by the Australian, state and territory governments — housed over 1000 items and produced 50 synthesis reports on the Council of Australian Governments' Closing the Gap commitments. The clearinghouse began operating in 2009, lost funding in 2014 and ceased publishing new synthesis reports in 2016, after which there was no whole-of-government approach to publishing and analysing evaluation reports for policies and programs affecting Aboriginal and Torres Strait Islander people. More detail on the roles and functions of the Closing the Gap Clearinghouse can be found in chapter 3 and appendix C.

Sources: ANAO (2019a); Australian Institute of Criminology and the Council of Attorneys-General (2016); Department of Education and Training (Vic) (2018); Department of Finance (1993, 1995).

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 the 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

involvement¹⁴ 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 Auditor-General, 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 Reporting Framework in 2009, a move that brought back program outcomes to performance reporting, but ultimately yielded no systematic 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 passing of 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 agencies can demonstrate the performance of their operations, while retaining flexibility in how this information is delivered. 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)

¹⁴ Executive may include the Secretary, Deputy Secretary and/or First Assistant Secretary or equivalent.

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 2.3). The framework underpins the PGPA Act (and its associated Rule).

Box 2.3 The Enhanced Commonwealth Performance Framework

The Enhanced Commonwealth Performance Framework came into force under the *Public Performance, Governance 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.

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).

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 2.3).

... 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 2.4). This function would:

- provide guidance and support for agencies on best-practice evaluation approaches
- drive a service-wide approach to evaluation and 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

The Review also proposed an evaluation profession be established within the APS.¹⁵

Box 2.4 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).

¹⁵ 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. The Review's implementation guidance points to heads of professions developing 'core competencies at different levels of the profession and learning and development frameworks' (p. 200).

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 about 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 DoF 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 the 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, the changes will establish central, whole-of-government functions and guidance material will be developed to support evaluation activity across the APS.

Other jurisdictions also adopt a centralised approach

Many 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 well established central agency oversight of evaluation and evaluation strategies (table 2.1). The Northern Territory Government has recently developed and released a program evaluation framework, 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.

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 2018* (US) (assented to by the President in January 2019), requires agency heads to appoint Evaluation Officers and to issue annual evaluation plans.
- 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)

Table 2.1 Central evaluation frameworks in 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 that 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 principles, methodologies (with examples), and a five-step evaluation process that evaluators should follow. It also provides templates and advice on planning an evaluation. The Program Evaluation Unit has dual roles of providing evaluation services and promoting evaluation as a core business function across the public sector. A sunset clause applies to any program over \$5 million for any one year. Programs subject to the 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>Program Evaluation Guidelines</i> support the application of evaluation within the context of government priorities. The guidelines provide high-level guidance on planning, commissioning, and managing evaluations as well as communicating evaluation findings and enhancing evaluation capability. They also provide links for further reading on topics covered.
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 the review required 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 the Department of 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.
Northern Territory	Department of Treasury and Finance	The first <i>Northern Territory Government Program Evaluation Framework</i> was released in May 2020. The principles-based framework outlines types of evaluation and evaluation's role in the program and budget processes, as well as guidance on evaluation scale and building evaluative capacity. The Framework's sunset clause applies to programs that impact the Territory's operating balance by \$1 million or more in any one year, and sets funding for no more than five years. Ongoing funding will be informed by evaluation.

Sources: ACT Chief Minister's Department (2010); Department of Treasury (WA) (2015); Department of Treasury and Finance (NT) (2020); Department of Treasury and Finance (VIC) (2019); DPC (NSW) (2016b, 2016a); Treasury (QLD) (2020).

Strategies and forward planning

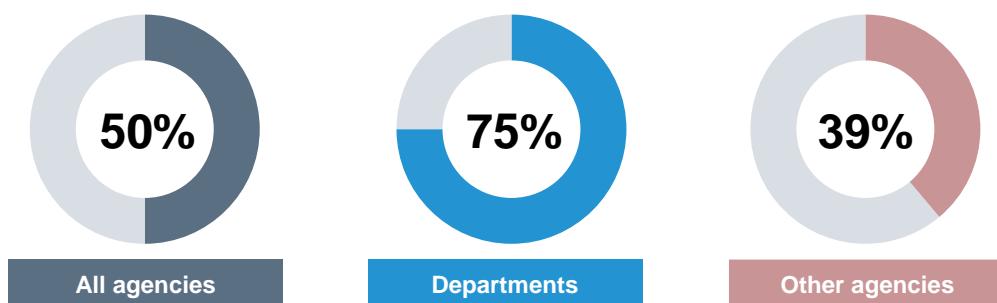
Evaluation strategies formally outline how agencies intend to approach evaluation. If well designed and followed, evaluation strategies can help agencies undertake evaluations in a consistent manner, and facilitate the strategic allocation of evaluation effort and resources to areas of greatest importance.

A number of Australian Government agencies have evaluation strategies or policies to guide how evaluation is planned for and undertaken within their agency. While half of all agencies responding to the Commission's information request said they had an evaluation policy or strategy in place, most departments (75 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, Transport, Cities 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 had evaluation strategies but they are not available on their websites.¹⁶

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



^a Based on responses from 52 agencies: 16 departments and 36 other agencies.

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

¹⁶ Strategies and policies were provided to the Commission (through its information request) before changes to Australian Government departments came into effect in February 2020. Analysis of strategies and policies reflects the situation as at December 2019, prior to the changes taking effect.

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 in chapter 6)
- require evaluation be considered when new programs are proposed
- provide guidance on how to develop and prepare a program logic.

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 is 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 Aboriginal 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

Department of Industry, Innovation, and Science	Department of Social Services	Department of Health	Department of Agriculture and Water Resources	Indigenous Advancement Strategy (NIAA)	Department of the Environment
Integrated Evaluation as core business, initiated at NPP, results communicated.	Integrated Conducted through life of intervention to inform future policy.		Integrated Initiated at program planning stage. Consistent measures.	Integrated Planning from outset. Past findings inform new. Learning, not compliance.	Integrated Planning from outset. Past findings inform new policy.
Fit-for-purpose Scale of resources, evaluation method.	Tailored Scale of resources, evaluation method.	Proportionality Scaled for size, risk.	Fit-for-purpose Scaled for size, risk.	Fit-for-purpose Appropriate to context, scale is proportional.	Fit-for-purpose Appropriate to context, scale is proportional.
Evidence-based Robust methods, careful data-gathering methods.	Robust Independent evaluation, best practice methods.	Quality Methodological rigor, oversight.	Evidence-based Robust methods, best data, baselines obtained.	Evidence-based Robust methods for real-world settings, appropriate data.	Evidence-based Robust methods, appropriate data, impact focused.
Independent Level of separation between evaluators and program areas.		Independence Check for influence. Separation between program and evaluation.	Independent Evaluation independent from program delivery area.	Independent Governance and evaluators separate from delivery.	Independent Evaluators should have some separation from delivery area.
Transparent Internal release expected ^b , work towards more public release.	Transparent Default is public release. Decisions and rationale recorded.	Transparency Stakeholders engaged, appropriate reporting.	Transparent Basis for demonstrating Department's accountability.	Transparent Publicly available with considered consent.	Transparent Results shared internally, made public to provide confidence.
Timely Guided by timing of critical decisions.	Useful Timely, relevant questions, relevant reporting.	Utility Relevant, useful, credible and timely.	Specific & timely Timing of critical decisions, time to adjust policy settings.	Timely Guided by timing of critical decisions.	Timely Guided by timing of critical decisions.
	Ethical Sensitivity to context. Use AES ^c guidelines.			Ethical Highest standards for Indigenous Australians.	
		Impartiality Objective consideration of findings.	Improving capability... through practice, learning from evaluation.	Impact-focused Testing causal explanations.	Outcomes-focused Are conditions improving?
		Consistency Conducted according to guidelines and requirements.		Respectful Collaborative, strengths-based. Integrates diverse perspectives.	

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

Sources: DAWR (2017); Department of the Environment (2015); DoH (2016); DIIS (2017); 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 include an ethical principle or provide guidance on how evaluation can be conducted ethically (figure 2.2). 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 (box 2.5). 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.

And specifically in the context of evaluation, 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 settings (Gibb et al. 2019). The Australian Evaluation Society (AES) also makes it a requirement for membership that evaluators adhere to their code of ethics (AES 2013a). 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.

Box 2.5 Ethical conduct in research guidelines

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) Code of Ethics for Aboriginal and Torres Strait Islander Research

The AIATSIS Code promotes ethical and responsible practice in research that impacts on or is of particular significance to Aboriginal and Torres Strait Islander people. It is structured around four principles, with each principle giving rise to several responsibilities:

1. Indigenous self-determination — including recognition and respect; engagement and collaboration; informed consent; and cultural capability and learning
2. Indigenous leadership — including Indigenous-led research; Indigenous perspectives and participation; and Indigenous knowledge and data
3. Impact and value — including benefit and reciprocity; and impact and risk
4. Sustainability and accountability — including Indigenous lands and waters; ongoing Indigenous governance; and reporting and compliance.

The National Health and Medical Research Council (NHMRC) ethical conduct in research with Aboriginal and Torres Strait Islander Peoples and communities

The National Health and Medical Research Council (NHMRC) considers that 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 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 (2020b); NHMRC (2018a).

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 Trade (DFAT 2020) and the former DIIS (DIIS nd). The NIAA also has an evaluation work plan that is endorsed by an Indigenous Evaluation Committee (DPMC 2018b, p. 10).

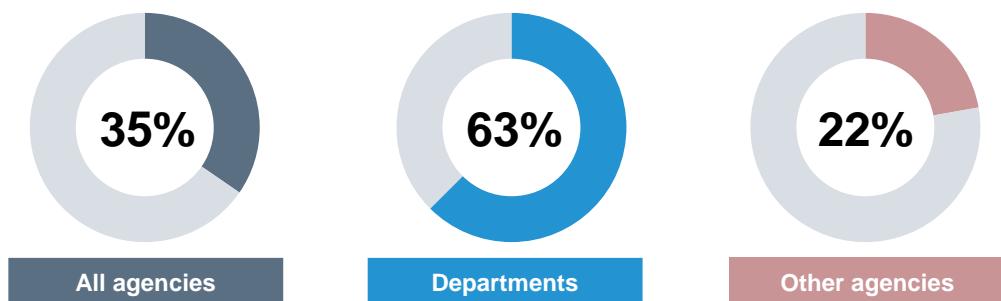
Publishing a forward plan promotes transparency, scrutiny and trust of an agency's operations and performance. Developing an evaluation plan also provides an opportunity for agencies to consider how they will evaluate programs, and what data, capabilities and resources are required to undertake evaluations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

Who undertakes evaluations?

About 35 per cent of Australian Government agencies and 63 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.

Figure 2.4 Agencies with a central evaluation unit^a



^a Based on responses from 52 agencies: 16 departments and 36 other agencies.

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

Most dedicated evaluation units:

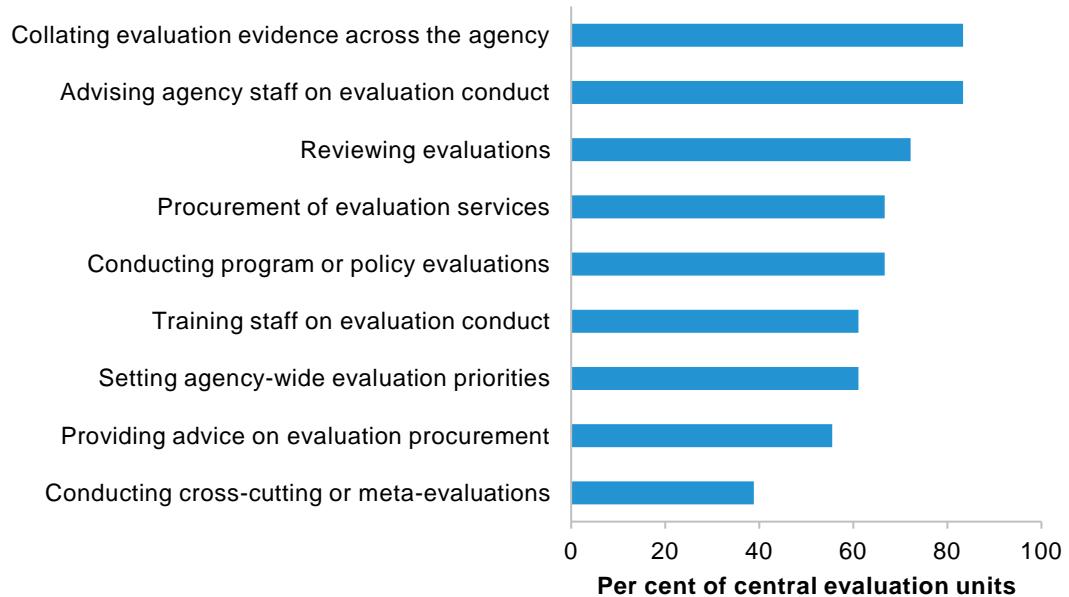
- advise agency staff on the conduct of evaluations
- procure evaluation services from external providers
- review evaluations conducted within the agency
- collate evaluation evidence from across the agency.

Responses to the Commission’s information request indicate that dedicated evaluation units have a range of roles. Two-thirds of evaluation units conduct evaluations (figure 2.5).

Resourcing for central evaluation units vary. For example, the evaluation unit within the former DIIS had 15 full-time equivalent employees while the central evaluation unit of another agency has an annual budget of about \$190 000.

It is common practice for Australian Government agencies to use external evaluators to evaluate policies and programs. In a background paper for the Independent Review of the APS, Bray, Gray and ‘t Hart (2019, p. 17) 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 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 18 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.¹⁷

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’ policies or programs considered to be of strategic importance, while evaluations of less important policies and programs were conducted by internal evaluation units.

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 set 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 capability, 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,

¹⁷ Tenders in a given year for Australian Government departments were filtered for the word ‘evaluation’ in the description. Further manual filtering removed tenders that were obviously not program or policy evaluation. Lack of detail provided in the tender descriptions however means that results are indicative only.

expertise to a program area, and developing capacity within the APS. The trend towards outsourcing has been growing, and continues to grow.

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 have an adverse effect on 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 burden 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 AES 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 former DIIS's Evaluation Strategy also states that evaluation findings will be published where appropriate (DIIS 2017, 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 published. Of the 509 evaluations identified, only 44 per cent were published. Publication rates vary between agencies with eight per cent of agencies reporting they 'always' published evaluation reports and about a third 'very often' while others publish relatively few (figure 2.6).

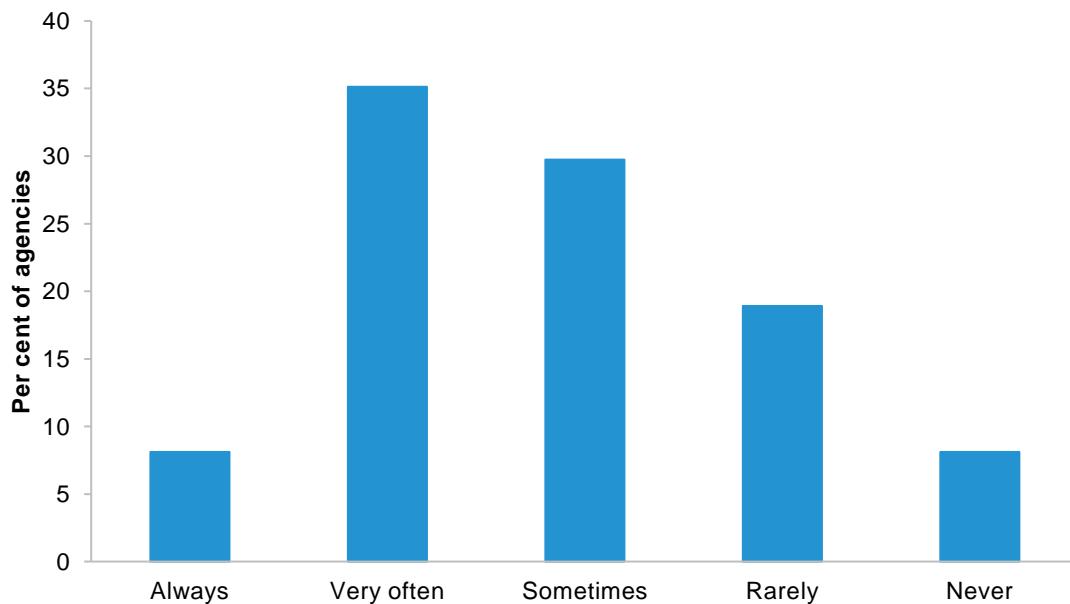
The Commission heard about a reluctance to publish evaluations when the findings were not positive and some suggested that requiring evaluations to be published could lead to pressure to remove or moderate findings. One participant also commented that:

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 37 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 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
- 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 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, as well as the potential to suspend parents' income support payments, 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 require (and publish) management responses to evaluation findings — examples include the former Department of Infrastructure, Transport, Cities and Regional Development and the former DIIS. Management responses provide a useful and transparent way to demonstrate how evaluation findings are influencing decision making and linked to the policy development and implementation process.

2.4 What lessons are there for the Indigenous Evaluation Strategy?

While many point to the importance of high-quality evaluation for producing evidence on the effectiveness (or otherwise) of policies and programs and how to improve outcomes for Aboriginal and Torres Strait Islander people (chapter 1), the history of evaluation in the APS is not exactly a success story. As described above, it is more a story of ebbs and flows.

When thinking about how to ensure an Indigenous Evaluation Strategy is effective in bringing about real change, it is important to ask: ‘what have we learnt from the past?’

Reflecting on past arrangements, a number of government agencies and other participants commented that one of the problems with previous approaches to evaluation that featured centralised requirements was that it became a ‘tick-a-box’ or compliance exercise rather than an exercise in valuing evidence.

While an evaluation system centred around compliance may mean more evaluation is being undertaken, the benefits of evaluation will be limited unless it is underpinned by a culture that values the knowledge generated by evaluation. For example, compulsory publication of evaluation reports may aid in transparency, but is of limited value if it leads an agency to be more risk averse and direct evaluation towards policies and programs where success is expected and little can be learned, with more marginal programs (where the benefits of evaluation would be greater) ignored. Wherever an evaluation rule exists, there could be incentives to do it cheaply or perfunctorily.

A good evaluative culture mitigates against these incentives. Chapter 7 discusses the importance of an evaluative culture and incentivising agencies to evaluate with the purpose of improving policy and program outcomes for Aboriginal and Torres Strait Islander people.

Previous evaluation approaches also demonstrate the risks of installing a blanket requirement that everything be evaluated. One of the more significant critiques of the evaluation strategy of the late 1980s and early 1990s was that the requirement to evaluate every program every three to five years resulted in a system that was cumbersome and that hindered the ability to direct evaluation effort to where it was most needed or where there was the greatest benefits (appendix C).

A structured and robust approach to setting evaluation priorities (both on a whole-of-government level and at an agency level) reduces the risk of similar problems occurring again. This has been front and centre in the Commission’s mind as it has designed an approach to setting evaluation priorities under the Indigenous Evaluation Strategy (chapter 6).

Finally, while many Australian Government department evaluation strategies share a common set of core principles (figure 2.2), the effective evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people requires principles that extend beyond those that agencies currently use. As such, even agencies with relatively well-developed evaluation strategies will need to change how they do evaluation if they are to implement the Strategy effectively.

A strategy that aligns with the new National Agreement on Closing the Gap (box 1.2) will be a useful tool to help agencies fulfil their new responsibilities under the Agreement. An Indigenous Evaluation Strategy that is implemented alongside the cultural shift initiated by the National Agreement’s third priority reform area¹⁸ stands a greater chance of creating real change and realising the Strategy’s ultimate goal of improving the lives of Aboriginal and Torres Strait Islander people.

¹⁸ Through the third priority reform area, the Parties to the National Agreement on Closing the Gap have committed to ‘systematic and structural transformation of mainstream government organisations to improve accountability and respond to the needs of Aboriginal and Torres Strait Islander people’ (JCOCTG 2020, p. 11).

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 a number of areas:
 - the lack of evaluation and the ad hoc way decisions are made about what to evaluate
 - 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 on 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 across the policy domains identified in the National Agreement on Closing the Gap found significant evidence gaps on the impact of policies and programs and how outcomes could be improved. That said, some themes from the evidence are that successful policies and programs are: implemented with an evidence base behind them; designed, developed and delivered with Aboriginal and Torres Strait Islander people and communities; and are flexible and adaptable to 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 a number of areas. They included concerns about: the extent and ad hoc nature of evaluation; the quality of evaluations; evaluation being an afterthought rather than part of the policy cycle; 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 thin). 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)

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)

And Lowitja Institute noted that:

Despite years of targeted reforms, we are still yet to see significant benefits to the lives of Aboriginal and Torres Strait Islander peoples. This is in part because of a lack comprehensive evaluation practice embedded throughout the Australian Public Service (APS) and into the policy cycles of Aboriginal and Torres Strait Islander affairs. (sub. DR157, p. 1)

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 found that just 8 per cent¹⁹ of Indigenous-specific programs had been evaluated (either during or after implementation) (Hudson 2016).
- 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 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 1249 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).

¹⁹ The study identified 1082 Indigenous-specific programs — 49 Australian Government programs, 236 state or territory programs and 797 programs delivered by non-government organisations (funded partially or fully 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)

For too long, we have seen crude tick and cross approaches as a prominent method of reviewing policies and programs in Aboriginal and Torres Strait Islander affairs, rather than thorough evaluation and careful consideration of findings to drive continuous quality improvement.
(Australian Medical Association, sub. DR116, p. 1)

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 said:

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)

A number of participants said that evaluations are often focused on the perspectives of government agencies rather than those of Aboriginal and Torres Strait Islander people and communities.

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. (AbSec – NSW Child, Family and Community Peak Aboriginal Corporation sub. 9, p. 4)

... the values and principles that inform evaluation design in relation to programs in Indigenous Communities — both in Australia and elsewhere — are not usually those of the Communities themselves, but represent those of the evaluators, mainstream organisations (including Government) responsible for the programs in question, and/or funding agencies. (National Centre of Indigenous Excellence, sub. DR158, p. 9)

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 agencies' 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.²⁰ Further information on the information request and more detailed results can be found in appendix B.

Fifty-two agencies responded to the information request, including 16 government departments,²¹ 32 other mainstream agencies, and four Indigenous-specific agencies. This

²⁰ 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.

²¹ The Department of Foreign Affairs and Trade is active in evaluating its programs and 57 of its evaluations were provided but is not included in the remaining analysis in this chapter. The Department's evaluations

was an overall response rate of 29 per cent. However, the response rate from departments (excluding the three parliamentary departments) was 94 per cent.

Agencies provided information on 307 evaluations conducted between 2016-17 and 2019-20. Nearly three-quarters of these 307 evaluations were published. Agencies also identified (but did not provide details on) a further 202 unpublished evaluations, giving a total of 509 evaluations (figure 3.1). Overall, 44 per cent of Australian Government evaluations identified by the Commission through the information request were published. Publication rates varied between agencies. Eight per cent of agencies said they ‘always’ published evaluation reports and about a third ‘very often’, while others published relatively few (appendix B).

Evaluation effort varies across agencies and topic areas

The responses to the information request show that evaluation effort varied across agencies — some agencies have undertaken 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 show that some (but not all) small agencies²² had been active evaluators in recent years, while some large agencies had undertaken very little evaluation.

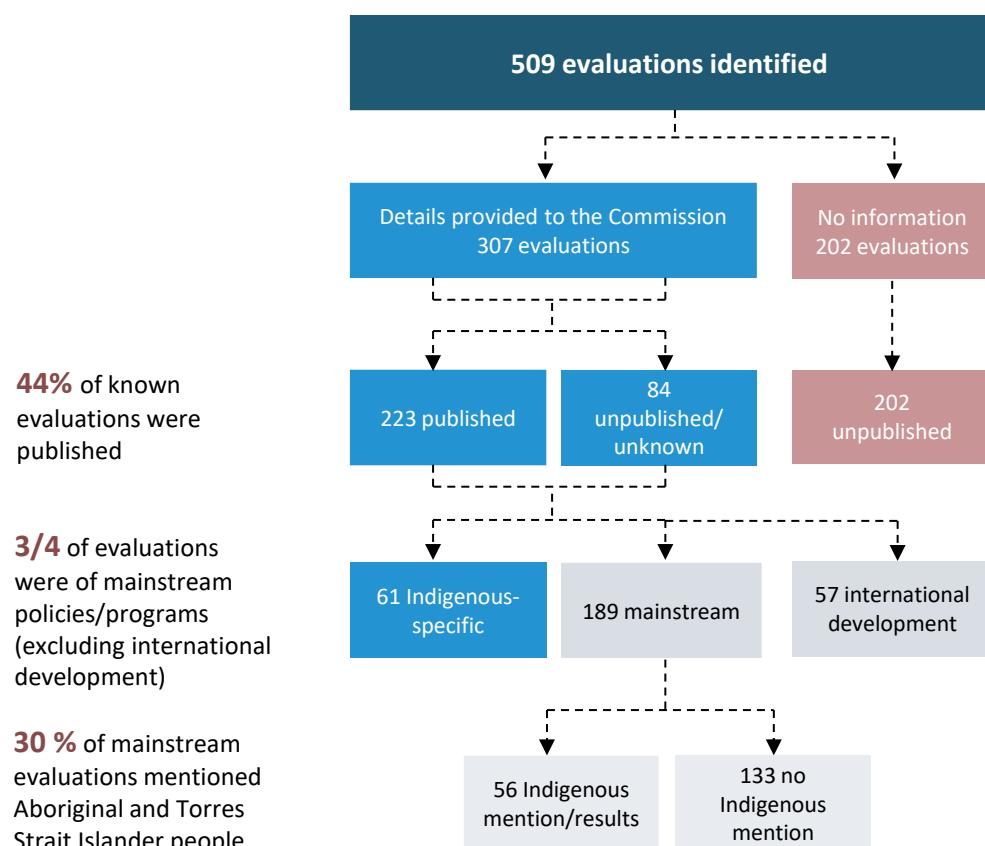
Some agencies were active evaluating a range of their mainstream policies and programs, however, only about 30 per cent of mainstream evaluations mentioned, or provided results on outcomes for, or effects of, mainstream programs on Aboriginal and Torres Strait Islander people.

The Commission looked specifically at evaluations provided by agencies that mentioned or provided results for Aboriginal and Torres Strait Islander people. Of these, there were 61 evaluations of Indigenous-specific policies or programs and 56 evaluations of mainstream policies or programs (figure 3.1). Some of these evaluations provided results on outcomes for Aboriginal and Torres Strait Islander people, while others provided more limited information (including, for example, the disaggregation of some data by Indigenous status or an acknowledgment of possible effects of policies and programs on Aboriginal and Torres Strait Islander people).

focus on international development and assistance programs that do not affect outcomes for Aboriginal and Torres Strait Islander people.

²² 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.

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



^a Counts represent information provided by the 52 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, programs and evaluations varied greatly, therefore, counts only provide an indicative picture of overall evaluation practice. Agencies provided details for 307 evaluations. A further 202 unpublished evaluations were identified by agencies but no information on these was provided.

Data source: Productivity Commission Indigenous Evaluation Strategy information request.

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).²³

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.

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 one-fifth 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. Only four evaluations of policies or programs mentioning or providing results for Aboriginal and Torres Strait Islander Australians included economic analysis. Chapter 9 provides more information on the use of data in evaluations (and how data collection can be improved).

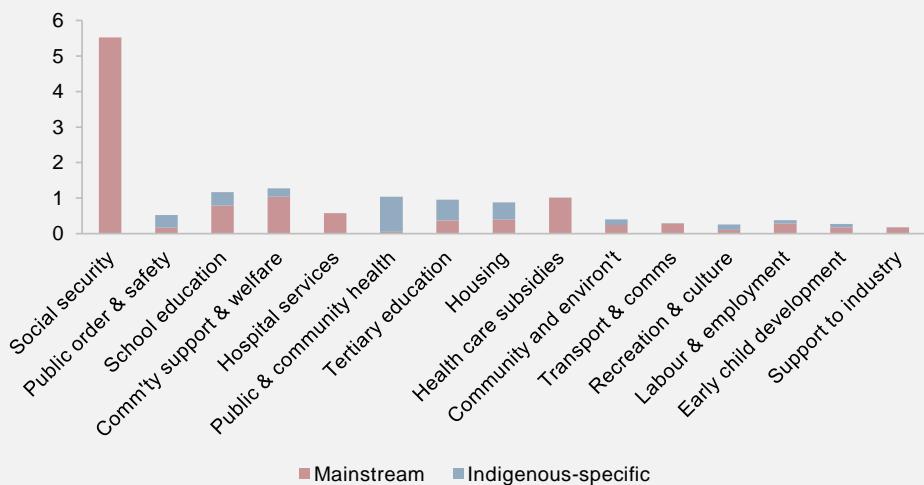
²³ There may have been evaluations in these areas (and others) that were not provided to the Commission, or evaluations undertaken by entities other than Australian Government agencies. Policies and programs may also have been evaluated in earlier years.

Box 3.1 Comparing patterns of evaluation and expenditure

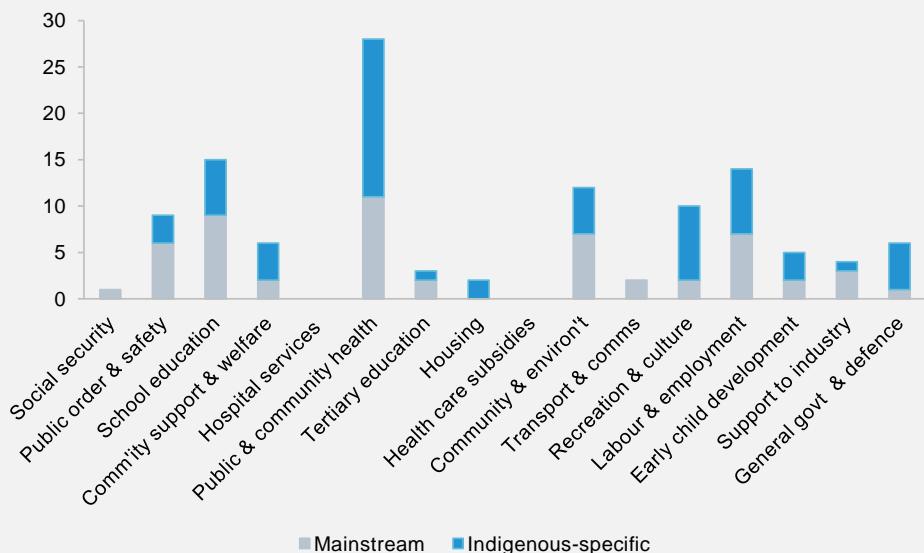
Data from the Indigenous Expenditure Report (SCRGSP 2017) show the distribution of Australian Government direct expenditure on Aboriginal and Torres Strait Islander people by service area. The areas attracting the most Australian Government direct Indigenous expenditure in 2015-16 (the most recent data 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)



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

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 30 per cent 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 56 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
- seventeen of 61 Indigenous-specific evaluation reports mentioned having Aboriginal and Torres Strait Islander people as part of the research or evaluation team.

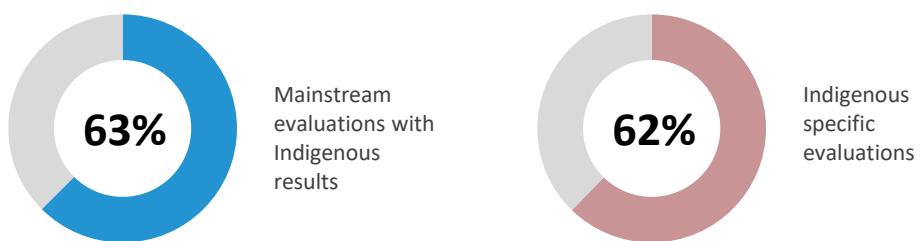
No mainstream evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people reported Aboriginal and Torres Strait Islander people being engaged in evaluation analysis or report authorship (although one did have an Aboriginal organisation as part of a reference group for the evaluation), or that the results were provided back to Aboriginal and Torres Strait Islander people. For Indigenous-specific evaluations about one-quarter engaged Aboriginal and Torres Strait Islander people in evaluation analysis or reporting and about 15 per cent 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.

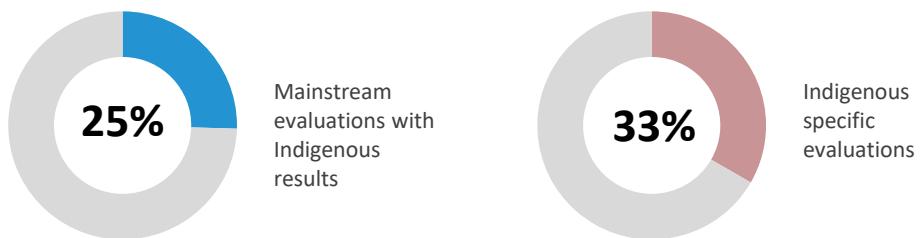
Figure 3.2 illustrates some characteristics of evaluations that mention or provide results for Aboriginal and Torres Strait Islander people.

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

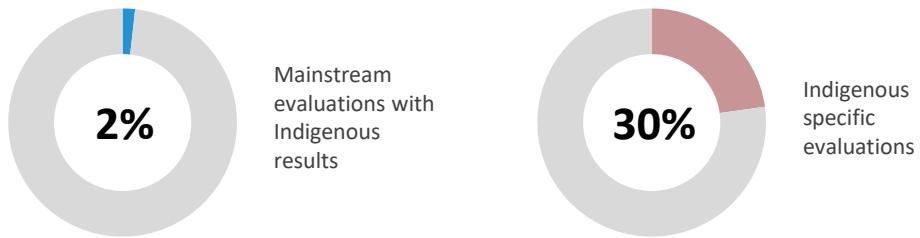
About six out of ten evaluations were done by external consultants



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 52 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, 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] 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. [Aboriginal Affairs NSW] will continue to work with the [NSW Coalition of Aboriginal Regional Alliances] 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-methods 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-methods 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. ... [Victorian Aboriginal Child Care Agency] 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 Volunteers²⁴ also reported that:

Over the past six years, [Indigenous Community Volunteers] 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 Aboriginal and Torres Strait Islander 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 (COAG) 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

²⁴ Indigenous Community Volunteers changed its name to Community First Development in early 2020.

-
- a general collection of research and evaluation on topics related to the Closing the Gap building blocks but not assessed by the Clearinghouse
 - 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 evaluations relating to both general populations as well as indigenous peoples (AIFS nd; AIHW 2017).

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 2017). The 36 Clearinghouse papers reviewing the evidence across a range of topics are available on the Australian Institute of Health and Welfare and Australian Institute of Family Studies websites (AIFS nd; AIHW 2017). 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 agencies²⁵ 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.

²⁵ 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 Trade²⁶ maintains an easy-to-use webpage providing links to a wide range of published evaluation reports
- links to published evaluation reports for the National Indigenous Australians Agency are available through published annual 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. However, as discussed earlier, after decades of governments developing policies and programs to improve the lives of Aboriginal and Torres Strait Islander people, the evidence about what works is thin.

What we do know, and it is unsurprising, is that programs and policies that have been informed by an evidence base are more likely to have successful outcomes (Bowes and Grace 2014, p. 3). This does not imply that policy makers should not experiment with new policy approaches, or address urgent policy issues, without evidence from previous evaluations. Rather, it means that a stronger evidence base improves the likelihood of a successful policy outcome.

There have been few recent attempts to synthesise learnings from evaluations of policies and programs that affect Aboriginal and Torres Strait Islander people. A 2018 literature review for the then Closing the Gap framework 10 Year Review noted that:

... apart from the 2012 review by the Closing the Gap Clearinghouse, there has been limited attempts to bring together and synthesise findings from research and evaluation across the various Closing the Gap initiatives to identify what has worked and what lessons have been learned from the implementation of the overarching Closing the Gap Policy framework. (CIRCA 2018, p. 13)

²⁶ The Department of Foreign Affairs and Trade has recently changed the administration of its evaluation functions but it continues to evaluate programs and evaluation reports remain available on its website.

The Closing the Gap Clearinghouse (when it operated between 2009 to 2014) produced three reports (over three years), which outlined high-level principles for successful programs in ‘overcoming Indigenous disadvantage’. Key reported principles for ‘what works’ included:

- flexibility in the design and delivery of policies and programs to take into account local needs and contexts. This requires understanding that issues are complex and contextual, the underlying social determinants, and the need to respect the effect of different cultures and languages (the importance of understanding context was also raised by a number of participants, box 3.3)
- community involvement and engagement in the development and delivery of programs, for example, through strong leadership, and community-member engagement through partnerships, networks and shared leadership
- the importance of building trust and relationships, including developing social capital to work collaboratively with stakeholders
- adequate resourcing, and planned and comprehensive policies and programs, including a well-trained and well-resourced workforce, with an emphasis on retention of staff
- continuity and coordination of services, including collaboration that builds bridges between public agencies and the community, and coordination between communities, non-government organisations and government to prevent duplication of effort (AIHW and AIFS 2013, p. 1; Al-Yaman and Higgins 2011, p. 2).

The Clearinghouse also identified a number of principles of ‘what does *not* work’ across the key building blocks of the then Closing the Gap policy framework:

- programs implemented in isolation: ‘one size fits all’ approaches, designed without context in mind (such as existing services, local culture and belief systems, and the physical, economic and social realities of the community)
- short-term funding and high staff turnover: insufficient time provided to build relationships between service providers and Aboriginal and Torres Strait Islander communities, and a failure to develop Aboriginal and Torres Strait Islander people’s capacity to provide services
- lack of cultural safety: programs designed without local Aboriginal and Torres Strait Islander community control and culturally appropriate adaptation, such as employing staff without an understanding of local culture, language and knowledges
- inflexible program delivery: where the materials delivered and the program structure is not accessible to the local community (for example, lack of transport may be a barrier for some people to access the program)
- external authorities imposing change and reporting requirements (AIHW and AIFS 2013, p. 5; Al-Yaman and Higgins 2011, p. 3).

Box 3.3 Participants highlighted the importance of putting ‘what works’ into context

Australian Council of TESOL Associations (ACTA):

... discussion about what works and does not work ... seems to be taken for granted that we all know what works. But the question should be: ‘what works for whom and according to what criteria?’ For example, with Cashless Debit Card you could say that this policy is ‘working’ for those who think it is a good idea. Likewise, the abolition of bilingual programs in the Northern Territory has ‘worked’ for policy makers who think that bilingualism holds kids back. But many language educators consider the termination of (good) bilingual programs to be a key factor in why the Closing the Gap targets have not been met. So what works for whom and why is a question that needs to be asked. (sub. DR179, p. 7)

BetterEvaluation:

Classical policy-level evaluation designs that aim to answer questions about ‘what works’ and then scale this up are ... [not useful] ... in conditions of volatility, uncertainty, complexity and ambiguity, or, where seeking to develop whole-of-government approaches to ‘wicked problems’ ... Instead, evaluation needs to be embedded in planning and implementation processes to ensure that delivery is adapted to local needs, opportunities and priorities, moving from finding out ‘what works’ to supporting local adaptation to make things work better in all contexts. (sub. DR144, p. 5)

Darling Downs and West Moreton Primary Health Network:

What works in one community is not going to work in another. It is unrealistic and disrespectful to suggest that such may occur for large programs across Australia. Program evaluation will need to be place based and published/promoted to recognize the places and localized results. (sub. 6, p. 3)

Jobs Australia:

The CDP [Community Development Program] program operates in regions where employment opportunities are often scarce, with the local community often facing high levels of adversity across a range of indicators which extend beyond the challenges associated with isolation and remoteness. While this is acknowledged in the Evaluation of Participation and Employment Outcomes for CDP, specific local factors are largely lost through the aggregation of the data and the limited criteria relating specifically to the various programmatic requirements such as participation, payment suspension and employment outcomes. (sub. 57, p. 6)

Many of these themes were also identified by the Queensland Productivity Commission in their report on *Service delivery in remote and discrete Indigenous Communities*. Services identified as working well, were those that:

- took care of root causes, rather than focusing on the symptoms
- adopted a developmental approach, including a strong sense of community ownership and control
- were people focused, and incorporated a ‘bottom-up’ approach to program design, decision making and service delivery that included community leadership and culture
- supported iterative learning and capacity building
- aligned with ‘place-based’ requirements, rather than jurisdictional, departmental or program boundaries. (QPC 2017, p. xiv)

Lessons (and knowledge gaps) in specific policy domains

This section outlines some lessons (and knowledge gaps) across the eight policy domains connected to the National Agreement on Closing the Gap's socio-economic outcome indicators and targets. It also covers what participants said (including participants' comments at the priority-setting workshop held by the Commission, chapter 6) about knowledge gaps and areas where further evaluation evidence could help support better policies and programs.

Safety (Families, children and youth)

On family violence prevention and early childhood programs delivered to Aboriginal and Torres Strait Islander communities, the Closing the Gap Clearinghouse found evaluation evidence to be limited. For example evaluations of mainstream early childhood programs often had small samples of Aboriginal and Torres Strait Islander children, and many evaluations focused on outputs, rather than outcomes (AIHW and AIFS 2013, p. 2). A report prepared for the Lowitja Institute also noted that while evidence was available for some programs delivered in the early years of childhood for Aboriginal and Torres Strait Islander children and families (such as maternal and child health services, early learning programs, and positive parenting programs):

The majority of evidence-based programs have not been tested through [randomised controlled trials] in Australia (let alone in Aboriginal communities or contexts), and even the best-supported programs have not necessarily been independently tested with multiple cohorts or populations, although there are a small number of exceptions to this. (Smith and Fox 2015, p. 31)

The available evidence for some of these programs supports the principles discussed above, including the importance of a program's responsiveness to local needs, particularly ensuring mainstream programs can be adapted to work effectively in Aboriginal and Torres Strait Islander communities. For example 'what works' on:

- *maternal and child health programs* are 'holistic programs that are responsive to the needs of target populations and local contexts' (shown to be effective in increasing antenatal health care and improving child health outcomes), as well as universal maternal and child health services and specialist support to 'address key determinants of child health such as maternal smoking or alcohol and drug use' (Fox and Emerson 2015, p. 17)
- *early learning programs* is 'attendance and participation in a quality preschool environment' (shown to be effective at improving short-, medium- and long-term health and development outcomes). Also a 'positive, engaging home learning environment is a strong predictor of good outcomes, and can help ameliorate the impacts of poverty and disadvantage' (Fox and Emerson 2015, p. 17)
- *positive parenting programs* is 'responsive and attuned parenting [which can] can significantly ameliorate the detrimental impacts of poverty and disadvantage, especially when combined with a rich home learning environment.' The efficacy of parenting

programs specifically on Aboriginal and Torres Strait Islander parents is less clear, in part because of challenges in adapting the cultural fit of programs based on Western norms of parenting and child rearing (Fox and Emerson 2015, p. 19).

According to Smith and Fox, the evidence on the benefits of investing in the early years is ‘compelling’, especially for children facing social and economic disadvantage (2015, p. 31). Some participants identified evaluations of policies and programs that build a strong start for kids (pre-natal to age 5) as a key priority. Maggie Walter, for example, noted:

The first 1000 days of a child’s life can be the most important. If funds are limited, this should be a priority for government policies and programs, as resilience is laid down by the age of four. (sub. 112, p. 2)

Others commented that evaluations in this area need to capture: the longer-term outcomes for children beyond the early years; the performance of the early childhood system (focusing on outcomes) rather than on participants in the system (focusing on inputs, such as enrolments); and the performance of Aboriginal and Torres Strait Islander community controlled services compared to mainstream providers in meeting the development needs of children. And on policies and programs aimed at addressing domestic and family violence, some argued for evaluation outcomes to be linked back to the *National Framework for Protecting Australia’s Children*.

Education

On school education, there are relatively few evaluations of policies or programs designed specifically for Aboriginal and Torres Strait Islander students (much of the literature covers overseas studies). Al-Yaman and Higgins (2011, p. 20), for example, noted in one review of 139 evaluations relating to early childhood and schooling, less than one in five of the programs and strategies were designed specifically for Aboriginal and Torres Strait Islander students. Based on the Commission’s analysis of recent evaluations of education policies and programs, most evaluations undertaken are of mainstream policies and programs (often at a fairly high level) and either do not mention results for Aboriginal and Torres Strait Islander people, or do so relatively briefly.

Many of these evaluations focus on improving school attendance but less focus is given to their effectiveness on outcomes. 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, p. 94).

Staff quality (in particular school leadership), parental and community engagement, and privileging local priorities (including the use and development of Aboriginal and Torres Strait Islander languages) are also factors highlighted as important to education outcomes, although evidence on how these can best be delivered across communities is thin.

Looking at high-performing remote schools, the What Works team from the National Curriculum Services found that many of the general principles discussed above on ‘what works’ applied to successes in remote schools (box 3.4).

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).

Participants identified a need to take a more holistic approach to the evaluation of the education system, starting with identifying what constitutes a good quality education (from the perspective of Aboriginal and Torres Strait Islander people), and what the education system can offer Aboriginal and Torres Strait Islander people. Also, that evaluations should focus on longer-term outcomes (experiences and success in higher education) rather than focus on inputs (attendance and enrolments). The role of languages in affecting student outcomes was also raised as an important factor that should be evaluated (discussed below).

Tertiary education was also identified as an area needing greater evaluation activity. Clair Anderson, for example, noted:

We consider that the evaluation of Indigenous higher education programs and policies, and respective educational outcomes, needs to be clearly identified as a key national evaluation priority. ... calls for a greater focus on evaluation in Indigenous higher education contexts has been repeatedly highlighted through national reviews, ministerial advisory councils, and empirical research. Yet, minimal tangible action by the Australian Government has occurred. (sub. 105, p. 5)

A recent report on *Strengthening Evaluation in Indigenous Higher Education Contexts in Australia* also concluded ‘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).

Similar gaps appear to exist in the vocational education and training sector. A report commissioned by the National Centre for Vocational Education Research noted that ‘we know’ vocational education and training and higher education lead to higher levels of employment, and:

Successful engagement is built on community ownership, genuine partnerships with communities, respect for cultural knowledge and local capabilities, integration of cultural knowledge into training, and alignment of education and training with aspirations and, in the case of remote areas, local employment opportunities. (Ackhurst, Polvere and Windley 2017, p. 1)

However, the report also noted many things that are not known from research in the VET sector, including:

- Why increasing training completion remains a challenge.
- Why increases in tertiary education levels have not led to improved employment outcomes.
- How the VET sector can adapt to better meet the needs of Indigenous communities, acknowledging and responding to their educational aspirations and local employment opportunities. (Ackhurst, Polvere and Windley 2017, p. 1)

The Alliance of First Nations Independent Education and Training Providers advocated for evaluation evidence to prioritise the effectiveness of Indigenous Registered Training Providers relative to mainstream services in terms of: completion rates and job readiness; the quality of online course service delivery; tackling social issues, such as rates of re-offending; and meeting cultural obligations (sub. DR141, pp. 4–5).

Justice, including youth justice

There have been several reviews (including Royal Commissions), looking at the impact of the justice system on Aboriginal and Torres Strait Islander people, which have identified broad principles of reform to reduce the incidence and impacts of involvement in the justice system. Evidence from evaluations of specific measures (such as diversion, early intervention, and tertiary programs) have also been undertaken, but on a more ad hoc basis.

The evidence from these reviews has established principles for successful justice programs in line with those for successful programs in other policy domains, including better engagement, Aboriginal and Torres Strait Islander leadership, cultural safety, and adequate resourcing. In addition, learnings from successful justice policies and programs focus on:

- effective collaboration (including collaborative case management) across organisations (government and non-government) and between Aboriginal and Torres Strait Islander and non-Indigenous individuals and communities

- holistic and comprehensive approaches, with agencies engaging in early intervention, diversion or restorative justice programs. These can help achieve positive outcomes for clients with multiple and complex needs that can lead to offending, for example, alcohol and substance misuse, unemployment and lack of education (Higgins and Davis 2014, p. 11). Box 3.5 provides an example of these key themes in an evaluation of the Victorian County Koori Court.

Box 3.5 Success factors identified in the County Koori Court (Victoria)

Internal factors

- Meaningful participation of Elders and community in the sentencing conversation, and adherence to the principles of the Koori Court model.
- Court orders should link the Accused to services or activities to address causes of offending behaviour that have been identified during the sentencing conversation.

External factors

- Insuring that local services are available to assist in the rehabilitation of the Accused such as through attendance in court or provision of services.
- The sentencing options available to the Court must be flexible enough to assist in the rehabilitative processes that are identified as appropriate for the Accused.

Source: Dawkins et al. (2011, p. 50).

A systems-wide approach reflects one of the key themes noted in the Royal Commission into Aboriginal Deaths in Custody (RCIADIC) that:

... the more fundamental causes for the over-representation of Aboriginal people in custody are not to be found in the criminal justice system but in those factors which bring Aboriginal people into conflict with the criminal justice system in the first place. ... the most significant contributing factor is the disadvantaged and unequal position in which Aboriginal people find themselves in the society-socially, economically and culturally. ... The report examines the position of Aboriginal people in relation to health, housing, education, employment and income; it discusses the land needs of Aboriginal people. It shows how the attitudes of the dominant non-Aboriginal society, racism both overt and hidden and institutional racism, adversely affect Aboriginal people. It shows how some laws bear unequally upon Aboriginal people. (RCIADIC 1991, pp. 1.7.1, 1.7.2)

Participants commenting on this policy domain focused on addressing systemic issues that underlie important impacts and outcomes for Aboriginal and Torres Strait Islander people — rather than on specific targets or outcome indicators. The National Justice Project, for example, advocated:

... that examining outcomes regarding systemic discrimination in the delivery of government programs and services, and in interactions with the justice system, should be of the highest priority in any Indigenous Evaluation Strategy. (sub. 51, p. 4)

Participants to the priority-setting workshop also highlighted a need to better understand the provision of holistic justice services, including diversion and prevention programs, as well

as wrap around services, such as youth legal services, bail support, family violence programs, and disability support. With particular interest in understanding the relative costs of these community/diversion programs compared to imprisonment.

Significant evidence gaps were also identified by participants on disability, recidivism, and on women and children in the justice system. Participants in the priority-setting workshop also focused on improving data gaps in over-policing, discrimination and racism experienced by Aboriginal and Torres Strait Islander people. On recidivism, a report by COAG also noted a ‘critical’ need to improve the evidence base about what works to improve outcomes for prisoners and ex-prisoners.

To be able to properly assess what works to improve Aboriginal and Torres Strait Islander prisoners’ employment outcomes and reduce recidivism, information that is not currently collected or reported about prisoners is needed. It is also critical to have more rigorous measurement and evaluation of the impact of programmes and services on ex-prisoners. (COAG 2016, p. 48)

Participants also considered it important for governments to respond to key completed evaluations, for example, of the Indigenous Legal Assistance Program, and Royal Commission inquiries, such as the Royal Commission into Aboriginal Deaths in Custody.

Health and wellbeing

Al-Yaman and Higgins (2011) found a relatively high proportion of both mainstream 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’ (AIHW and AIFS 2013, p. 4). Bowes and Grace 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’ (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. Participants to this study identified the need for greater evaluation evidence focused on how mainstream programs deliver benefits for Aboriginal and Torres Strait Islander people. Some examples included: the National Disability Insurance Scheme (AMSANT, sub. 81, p. 6); mainstream aged care services (AMSANT, sub. 81, p. 6); the Medicare Benefits Schedule billing for medical services, and access to Pharmaceutical Benefits Scheme medicines to Aboriginal communities (Central Australian Aboriginal Congress, sub. 48, pp. 7–8); and the tax rebate for private health insurance. (Participants also highlighted a number of other mainstream and Indigenous-specific policies and programs that could benefit from greater evaluation evidence, box 3.6.)

A study by Eades et al. lamented the lack of meaningful research and evaluation looking at the actual 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)

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 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).

Participants to the priority-setting workshop also argued the need for evaluation to focus on holistic approaches to Aboriginal and Torres Strait Islander health, reflecting on the wider social and cultural determinants of health and wellbeing being developed across some communities. This was emphasised in the areas of mental health and social and emotional wellbeing, as well as disability, which many participants considered should be priorities for evaluation in their own right. The Royal Australian and New Zealand College Of Psychiatrists, for example noted:

This overarching target and the specific targets across all areas (Families, children and youth; health; education; economic development; housing; justice, including youth justice; land and water), can only be effectively addressed if the mental health needs of individuals and communities are met through appropriate resourcing. Research shows that the social and emotional wellbeing of Aboriginal and Torres Strait Islander peoples is underpinned through receiving specialist support to address mental health concerns related to intergenerational trauma, racism and ongoing marginalisation and discrimination. (sub. DR138, p. 6)

Box 3.6 Health and wellbeing evaluation priority areas identified by participants

- 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 and social and emotional wellbeing (Beyond Blue, sub. DR121, p. 1; IUIH, sub. 53, p. 12; RANZCP, sub. DR138, p. 7).
- Framework Agreements in Aboriginal health; state- and territory-based health planning forums; access to Pharmaceutical Benefits Scheme medicines in remote areas; access to Medicare Benefits Schedule billing for services to Aboriginal communities; Aboriginal health provision within the Department of Health; whole of population public health strategies, such as 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; RANZCP, sub. DR138, p. 7).
- Mainstream programs such as the NDIS/aged care for Aboriginal people (AMSANT, sub. 81, p. 6).
- 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).
- The education pipeline for Aboriginal and Torres Strait Islander health care workers (a suggestion put forward in the priority setting workshop).

Employment (Economic development)

Evaluations of impacts on Aboriginal and Torres Strait Islander people under this policy domain are relatively scarce. In 2014, a Clearinghouse resource sheet on *Success factors for Indigenous entrepreneurs and community-based enterprises* relied on ‘... a combination of case studies, literature reviews, program descriptions and analyses, internal organisational evaluations and research institute publications’, given the dearth of ‘rigorous evaluations of government or other programs that were aimed at improving Indigenous economic development’ (Morley 2014, p. 3).

While formative evaluations have been commissioned on key Indigenous-specific economic development and employment policies identified by participants to this study (namely the Community Development Program (DPMC 2019a) and Indigenous Procurement Policy (DPMC 2019f)), continued evaluation of these programs was still considered a priority (AMSANT, sub. 81, p. 5; Dreise et al., sub. 33, p. 2).

Participants at the priority-setting workshop also identified the following areas of policy that could also benefit from greater evaluation evidence.

- The (long- and short-term) impacts of COVID-19 pandemic policy responses on Aboriginal and Torres Strait Islander people. This could include: the employment impact of labour market measures (such as JobKeeper and JobSeeker support), and access to and outcomes of broader economic recovery measures (such as stimulus measures in the tourism or arts sector, or for production on the land).
- The impact of other stimulus and recovery measures, including from recent bushfires and floods, on Aboriginal and Torres Strait Islander people, measuring how Aboriginal and Torres Strait Islander people benefit from the recovery phase.

First Nations Media argued the need for greater evidence on ‘telecommunications infrastructure, internet and mobile phone access and digital literacy rates’, as ‘digital inclusion and digital literacy is an emerging need that now underpins equitable access to services for Aboriginal and Torres Strait Islander people’ (sub. 30, p. 11).

And the Alliance of First Nations Independent Education and Training Providers suggested that a number of priorities were missing under the National Agreement on Closing the Gap that warrant greater priority. They included adult Aboriginal and Torres Strait Islander learners (to obtain skills in language, literacy, numeracy and digital literacy), skills and courses most likely to lead to meaningful employment and improved wellbeing in the region in which they live, pathways off welfare into employment or self-employment, and support for Indigenous businesses and local communities (sub. DR141, p. 4).

Land and waters

There are several significant programs that facilitate or empower Aboriginal and Torres Strait Islander people to manage land and waters. One such program is the Indigenous Protected Areas program, where areas of land and water are managed by Aboriginal and Torres Strait Islander groups as protected conservation areas through voluntary agreements with the Australian Government. The Australian Government also funds Indigenous Ranger projects (formally known as the Working on Country program). Indigenous Ranger projects aim to provide meaningful employment, training and career pathways for Aboriginal and Torres Strait Islander people in land and sea management.

The outcomes from Indigenous Protected Areas and Indigenous Ranger projects have been examined relatively frequently, and evaluations and reviews consistently point to positive impacts for Aboriginal and Torres Strait Islander people, and the Australian community. Evaluations have identified economic, social and environmental benefits to the programs. Hill et al. (2013, p. 1) — through a literature review and analysis of spatial and quantitative data — collated evidence about the status, benefits and drivers of Indigenous land management (box 3.7).

Participants to the priority-setting workshop argued for greater evaluation evidence around land management, particularly in light of the recent bushfires in Australia, and the greater calls to adopt Aboriginal and Torres Strait Islander approaches to managing land and fuel loads through the use of traditional ecological knowledge.

The other key barrier to effective Indigenous land management identified by Hill et al. is the limitations of native title:

A key barrier experienced by many Indigenous people is lack of access to their traditional lands that are held under other forms of tenure. Although native title and land rights legislation are driving an increase in land ownership and access, regaining land is a long and difficult struggle for many. For others who have their ownership recognised, accessing vast areas of traditional lands in remote locations with very low rates of human occupancy is very challenging. (2013, p. 2)

While aspects of native title and the Native Title Act have been subject to reviews, including by the Australian Law Reform Commission (2015), and in a report to COAG (as part of a wider investigation into Indigenous land administration and use (Senior Officers Working Group 2015)), participants identified this as an area that would benefit from evaluation. In particular, evaluation evidence that focuses on the effectiveness of the native title system to allow Aboriginal and Torres Strait Islander people to use their land and water for economic, cultural and heritage benefits.

Housing

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 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.

The evidence that is available shows clear associations between housing and health, with housing affecting 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).

Box 3.7 Success factors and barriers to best practice in Indigenous Land Management

Hill et al. identified a number of factors support success in Indigenous Land Management (ILM).

- **Aboriginal and Torres Strait Islander motivation:** ILM involves an expression of identity, family linkages, customary law and responsibilities and obligations, which triggers a high level of Indigenous motivation to do land management.
- **Aboriginal and Torres Strait Islander governance:** This is most effective for ILM where: it is started by Aboriginal and Torres Strait Islander people themselves, traditional leaders are empowered; local views are taken into account; and agencies engage through supportive and facilitating approaches.
- **Hybrid economies:** Many successful ILM activities are based on commodities and practices that: i) can be sold in markets; ii) are underpinned by Aboriginal and Torres Strait Islander customs and iii) are supported to some extent by government investment.
- **Indigenous-specific government programs:** Indigenous-specific funding enables funding to be accessed based on cultural knowledge and practices.
- **Brokers and brokering organisations:** Successful ILM is helped by brokers who link Aboriginal and Torres Strait Islander organisations to the resources needed for ILM.
- **Relationships of trust, respect and mutuality:** Trust, respect and mutual interest underpin successful ILM. Time on Country is an important part of relationship building.
- **Diverse multimedia approaches for Aboriginal and Torres Strait Islander knowledge:** A range of multimedia approaches support Indigenous knowledge, while ensuring intergenerational knowledge transfer.
- **Collaborative two-way knowledge engagement:** Equitable engagement between Aboriginal and Torres Strait Islander and scientific 'tool boxes' for management facilitates successful ILM. Addressing underlying power imbalances and rights recognition is key if equitable engagement is to be achieved.
- **Aboriginal and Torres Strait Islander-driven planning:** This provides a way for Aboriginal and Torres Strait Islander people to control and shape their future, and to give it meaning.

A range of barriers can impede successful or best-practice ILM.

- **Limited respect, recognition and practical support for Aboriginal and Torres Strait Islander knowledges and worldviews:** The loss of traditional knowledge and language was a threat in many ILM situations
- **Limitations of native title rights recognition and access to traditional lands and waters:** Many Aboriginal and Torres Strait Islander people lack access to their traditional lands.
- **Limited access to resources for ILM:** Demand for ILM resources substantially outstrips supply.
- **Constraints in organisational and institutional capacity:** Many ILM organisations are fragile, under-resourced or do not have access to effective long-term administration, governance and infrastructure supports or systems.
- **Socio-economic and educational disadvantage faced by Aboriginal and Torres Strait Islander people:** Health and wellbeing issues affect the ability of Aboriginal and Torres Strait Islander people to undertake ILM.

Source: Hill et al. (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)

Cultures and Languages

This policy domain, as a standalone policy area, is a new feature in the Closing the Gap framework. Aboriginal and Torres Strait Islander cultures and languages, however, have always been a key principle underlying ‘what works’ across all other policy domains. For example, the key principles of leadership and shared decision making by Aboriginal and Torres Strait Islander people, and taking a bottom-up approach to engagement and policy and program design and delivery, ensures that policies and programs reflect a genuine understanding of Aboriginal and Torres Strait Islander people’s cultures and languages.

In the priority-setting workshop, participants noted government evaluations often do not know how to deal with culture, and how to measure cultural safety outcomes, despite culture being fundamental to services provided by Aboriginal and Torres Strait Islander people and organisations. A cultural lens, which builds on existing Aboriginal and Torres Strait Islander expertise and capability for designing community-led evaluations, is needed to ensure that culture is defined by Aboriginal and Torres Strait Islander people. And that stories, yarns and strength-based (rather than deficit-based) measures are the focus of evaluation design and results. Following this approach, participants noted that more evidence could be gathered, for example, on:

- the performance of frontline workers, their skills, and expertise to provide culturally safe services
- accessing/using mainstream services due to cultural or linguistic appropriateness of services, such as how a lack of interpreter services can prevent a just outcome in the justice system, or how NAPLAN, which fails to account for language proficiency, can impact educational outcomes

-
- bilingual teaching/learning approaches for Aboriginal and Torres Strait Islander students, especially in regional and remote locations (ACTA, sub. DR179, p. 15).

The exploration of languages in evaluations appears also to be an area where greater attention could be beneficial. The Australian Council of TESOL Associations (ACTA), for example, referred to an unpublished research paper, which examined twenty publicly available frameworks for, and evaluations of, Australian Government mainstream and Indigenous-specific programs over the last fifteen years. The paper found that of the few that considered the impact of languages:

... the main focus was on data collection, that is, with reference to its instrumental utility to evaluators ... [rather than on how] ... programs impacted on Indigenous language use – or how Indigenous language uses impacted on programs ... In other words, these evaluations paid no attention to how programs and evaluations related to, furthered, undermined or hindered the language-related rights just described. (sub. 87, p. 9 which draws on Blackwell (2019))

ACTA noted that language should be thought of not only in terms of a problem, barrier or impediment, such as if people cannot speak English or are illiterate, but rather as an asset, or a strength:

Policies need to take account of the way that languages actually work in society and for individuals, just as doctors need to understand how the heart works. Policy evaluations need to focus on whether, to what extent and how policies have understood and responded to how language(s) work, and what can and should be done to build on language strengths and language assets. That goes far beyond saying ‘People don’t understand English, therefore the policy is not working.’ (sub. DR179, p. 7)

4 Evaluation types, approaches and methods

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 and approaches 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 people. 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 a policy or program 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. 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 (table 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.

Table 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 or value-for-money evaluations.
Evaluation approach	The approach used to answer evaluation questions. For example, an impact evaluation could use an experimental, quasi-experimental or non-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, propensity score matching or case studies).
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 program recipients 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. Some potentially good policies and programs can fail to achieve their potential because of the way they are implemented. Making adjustments early and throughout the life of a policy or program can improve outcomes.

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 4.1. 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? Answering impact evaluation questions requires good data and statistical analysis to test what changes have occurred as a result of the policy or program.

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.

Box 4.1 **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. It uses a sporting platform to encourage participation. High-profile Aboriginal rugby league players from local communities 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 progress against 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 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, in the longer term, further evaluations should 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).

Experimental and quasi-experimental approaches infer the impact of policy or program through statistical comparison of a group or time period unaffected by the policy or programs. 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 4.2). 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.²⁷ 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 4.3.

RCTs, however, are not suited to evaluating some policies and programs. For example, it may not be ethical or possible to randomly assign some people in the same community to receive services and to allocate others to a control group (section 4.2). RCTs also depend on the principle of ‘all other things being equal’, which in practice can be difficult to achieve when policies are being adapted as they are implemented (Jowell 2003, p. 19).

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. Quasi-experimental methods vary in the extent to which they approach random assignment, with some sophisticated approaches to matching treatment and non-treatment groups, such as propensity score matching (box 4.4).

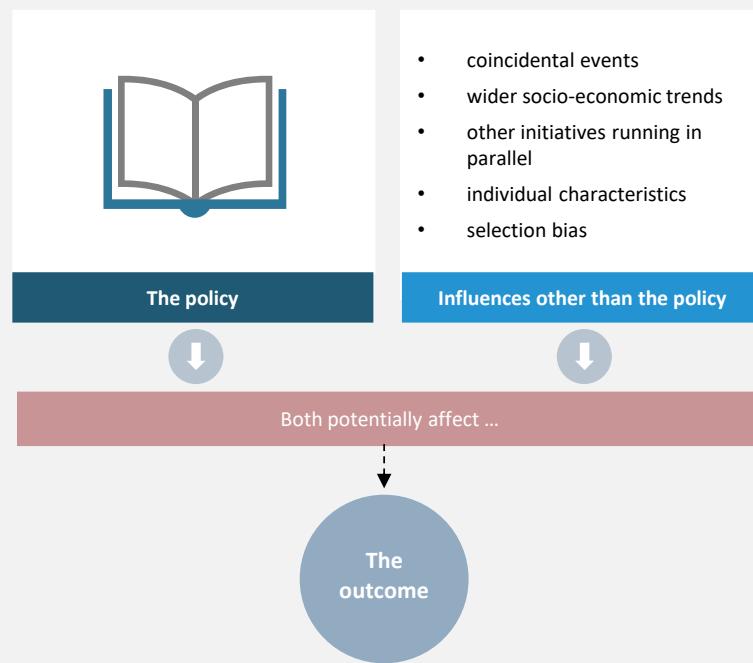
Experimental and quasi-experimental approaches can assess the net impact of a policy or program, however, they do not tell you how any observed change came about or if in another context or different scale the policy or program would produce the same outcome.

Theory-based approaches to impact evaluation can help shed light on how and why observed results occurred (White 2009, p. 3). They draw conclusions about impact by testing whether the causal chains that are thought to bring about change are supported by sufficiently strong evidence (and alternatives can be ruled out). Theory-based evaluation, using an explicit theory of change ‘tries to get inside the black-box of what happens between inputs and outcomes, and how that is affected by wider contexts’ (HM Treasury 2020c, p. 36).

²⁷ 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 4.2 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 (for example 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 (for example a new company establishes itself, increasing local labour demand).
- *Maturation*: where results reflect naturally occurring changes because of the passage of time (for example participants become tired, older, or their skills improve).
- *Testing*: where the administration of a pre-test affects post-test results (that is, 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 (for example 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 4.3 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).

A theory-based approach is most suited where: policy makers want to understand results in different places or contexts; where a suitable counterfactual cannot be developed; where a policy is being adapted and outcomes are not known and/or the policy is one of many within a complex system (HM Treasury 2020c, p. 37). Theory-based approaches include:

- systems mapping and modelling — used to generate, progress and test the theory of change through an iterative process of developing and testing a formal model of the system
- generative causation approaches, such as realist evaluation and contribution analysis, that seek to articulate underlying mechanisms or processes of change, and test the theory empirically to investigate whether, why or how the policy causes or contributes to observed results, and how context influences these (HM Treasury 2020a).

Box 4.5 provides an example of a realist evaluation of the Remote School Attendance Strategy.

Box 4.4 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.

<i>Method</i>	<i>Description</i>	<i>Pros and cons</i>
Simple pre-post	Compares participants' outcomes before and after an initiative is implemented (that is, 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 (that is, 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 (for example to gain access to the initiative).

(continued next page)

Box 4.4 (continued)

<i>Method</i>	<i>Description</i>	<i>Pros and cons</i>
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 (for example 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 (2020c, p. 76).

Box 4.5 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 Australian and state and territory 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 a one-size-fits-all approach is unlikely to be effective in improving school attendance in remote communities. 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).

There are also a number of other emerging approaches that are particularly useful for evaluating in complexity and supporting learning and adaptation, including developmental, participatory and empowerment evaluation (box 4.6).

Box 4.6 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 with these new insights.

Participatory evaluation

Participatory evaluation approaches involve collaboration between the evaluator and stakeholders. The evaluator and the stakeholders jointly make decisions about the evaluation questions, how the evaluation is conducted, and how evaluation results are shared. Participatory approaches can be used 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.

(continued next page)

Box 4.6 (continued)

Empowerment evaluation

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. It seeks to create positive change, as well as evaluating program outcomes.

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' (AIFS 2015).

One of the criticisms of empowerment evaluation is that the process is resource- and time- intensive for communities and service delivery organisations. Another is that empowerment evaluation can be more a method of teaching people about evaluation than an evaluation itself.

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)

Sources: AIFS (2015); Fetterman (2005), Patton (1994, 2016, pp. 526–527); Patton, McKegg and Wehiipeihana (2015); Pawson and Tilley (2004); Westhrop (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 always known but can be discovered through a process of trying things (Gamble 2008). Developmental evaluation is in some ways akin to an interactive and experimental prototype development process — design a policy or program, trial it, evaluate it, modify it guided by the evaluation results, then trial and evaluate a second version, and so on.

Both developmental and realist approaches can be useful in complex situations, including 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).

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 (BetterEvaluation 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 4.7).

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; Nous Group and CAYLUS 2017; 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 Marangka 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 by directing funds away from crisis response programs, adult prison and youth detention, and reinvesting it toward education, support and prevention.

New South Wales Treasury has also recently established an Aboriginal Economic Wellbeing Branch that is responsible for developing and rolling out an evidence-based and effective framework for including Aboriginal cultural values in cost-benefit analysis and policy evaluation more broadly (sub. DR180, p. 1).

As with impact evaluation, good quality economic evaluation relies on good quality data and evidence. Data on the costs of implementing policies and programs are often available, but estimating the benefits requires the same sorts of data, evidence and research methods as are needed for other forms of impact and outcome evaluation. Economic evaluation builds on evidence of outcomes to quantify those outcomes in monetary terms.

Box 4.7 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.

Figure 1: Cost-effectiveness analysis

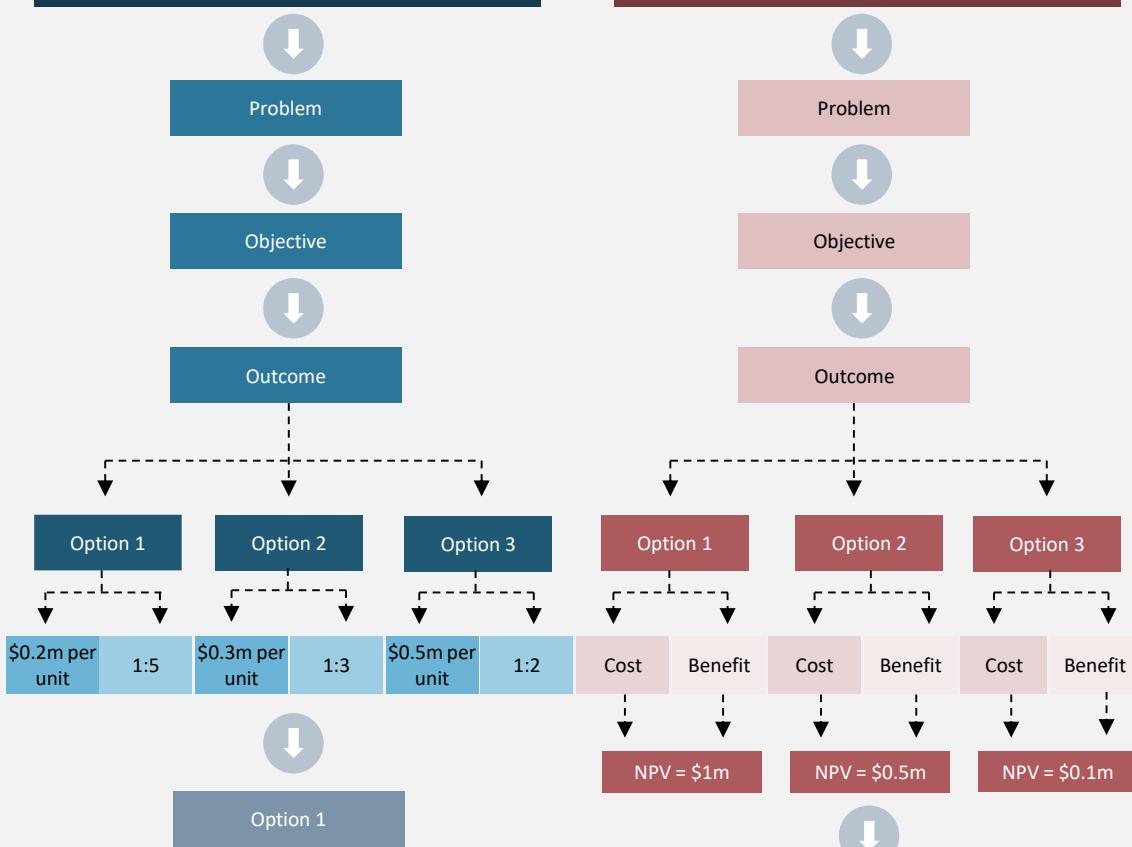
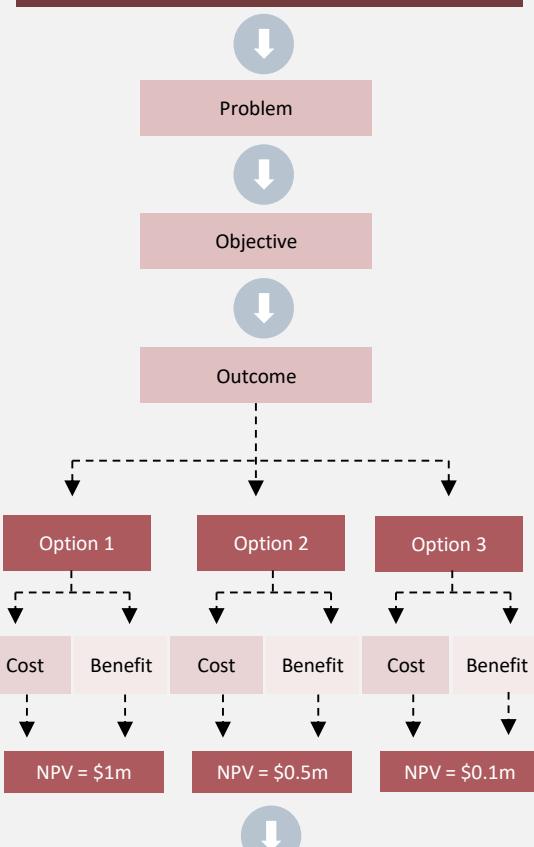


Figure 2: Cost-benefit analysis



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

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.

Commenting on Indigenous research methodologies, and noting that they had existed for thousands of years, McGregor, Restoule and Johnston (2018, p. 4) said they 'reflect how knowledge is understood and sought in the context of the worldviews, ontologies, and epistemologies of diverse Indigenous nations'. Dudgeon et. al, remarking on the recent increase in Indigenous research methods and methodologies, said it has 'shifted research beyond an acknowledgement of, or inclusion of Indigenous ways of knowing, towards a self-determining research praxis' (2020, p. 7).

A literature review on Indigenous research methods and methodologies found that they have the following key components:

- contextual reflection, in that the researchers or evaluators must situate themselves and the Indigenous Peoples with whom they are collaborating in the process
- inclusion of Indigenous Peoples in the research process in a way that is respectful and reciprocal as well as decolonizing and preserves self-determination
- prioritization of Indigenous ways of knowing. (Drawson, Toombs and Mushquash 2017, p. 17)

The review also found that the key distinction between Western and Indigenous research methods resides in research purpose and process — while Western approaches involve selecting 'an appropriate method to explore with as much scientific rigour as possible', Indigenous research 'done in collaboration with Indigenous people cannot only reveal knowledge, but also decolonize, rebalance power, and provide healing' (Drawson, Toombs and Mushquash 2017, p. 12). Indigenous methods are also culturally relevant, going beyond data collection to create relationships and support autonomy — they require that 'all components in the process embody the values of the Indigenous group involved' (Drawson, Toombs and Mushquash 2017, p. 15).

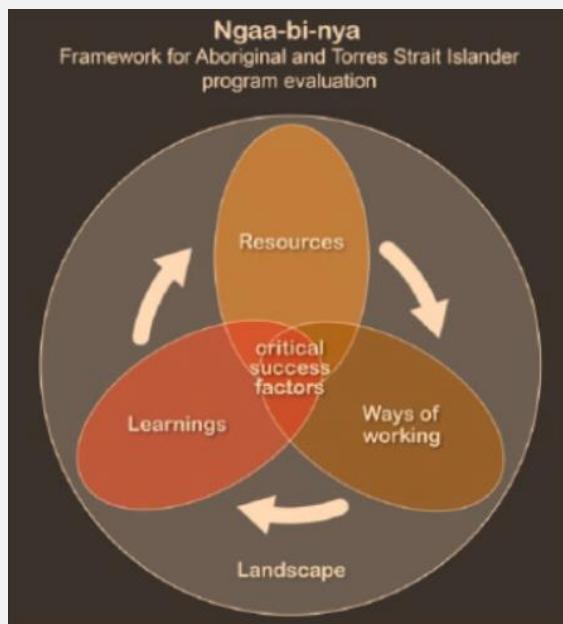
An example of an Indigenous approach is Ngaa-bi-nya — a program evaluation framework developed by Aboriginal academic and health researcher, Megan Williams, in collaboration with Aboriginal and Torres Strait Islander program providers and staff (Williams 2018). Ngaa-bi-nya provides a structure to analyse 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 4.8).²⁸

²⁸ 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) developed for the Lowitja Institute and University of Melbourne.

Box 4.8 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 and 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 learned. 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).

In New Zealand, there is a growing emergence and use of Māori evaluation theories, such as Kaupapa Māori (which means 'Māori way': Smith 1999; Wehipeihana 2019a). Kaupapa Māori evaluation can be characterised broadly as evaluation that:

- is controlled and owned by Māori
- is conducted for the benefit of Māori (although it may also benefit others)

-
- is 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); Moewaka Barnes (2000), 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)

Similarly, George Denny-Smith, Megan Williams and Martin Loosemore pointed out that:

... Indigenous perspectives and experiences have thus far been marginalised in policy evaluation. It is commonly people in a position of power who undertake evaluations and determine what should or should not be measured, and this can omit things that the people affected by evaluations see as being valuable. (sub. DR153, p. 5)

The Strategy's overarching principle recognises that for evaluation findings to be relevant and useful, Australian government agencies need to draw on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people when evaluating policies and programs (including when making decisions about what to evaluate, evaluation questions, approaches, methods and practices, and the interpretation of evaluation findings).

The Institute for Human Security and Social Change noted that:

... evaluation which is cross-cultural and where different worldviews and values are likely to encounter each other, needs to carefully consider the analysis processes, which build dialogue and connection between different worldviews. While this may not be a simple process ... if handled well it is likely to lead to more active attention to solving problems and working together for improvement. (sub. DR119, p. 3)

Collecting evidence — qualitative and quantitative methods

For each evaluation approach, there are a variety of data sources and methods of collecting data that can be used. Each has strengths and weaknesses (table 4.2, box 4.9). Data can also take many forms, including stories or pictures. Evaluations may be able to use existing data, but often new data needs to be collected. Regardless of the method used to collect data, it is important that the evaluation is based on sound evidence.

Table 4.2 Strengths and weakness of quantitative and qualitative data

	Strengths	Weaknesses
Quantitative	<ul style="list-style-type: none">• 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	<ul style="list-style-type: none">• 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
Qualitative	<ul style="list-style-type: none">• 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	<ul style="list-style-type: none">• Not always generalisable• Different perspectives may confuse the 'takeaway' message or conclusion• May be more subjective• Harder to assess validity without careful documentation

Source: adapted from Bamberger (2012a, 2012b).

Quantitative methods (such as surveys and questionnaires) have the advantage of being able to collect a lot of information in a short period of time. However, it is important that questions are 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)

For quantitative methods, the data must be of sufficient quality to represent the population affected by the policy or program. One of the challenges for using quantitative methods in the context of evaluating policies and programs affecting Aboriginal and Torres Strait Islander people is small sample sizes. For example, Aboriginal and Torres Strait Islander people may be only a small proportion of participants affected by a policy or program and this may make it difficult to produce statistically meaningful results (sampling is discussed further in chapters 9).

Qualitative methods, such as interviews, focus groups and case studies, can provide in-depth insights into people's experiences and perspectives (box 4.9). They have a particular strength in addressing questions of why and how (WHO 2013, p. 34). As Patton explained:

If you want to know how well children can read, give them a reading test (quantitative data). If you want to know what reading means to them, you have to talk with them (qualitative data). Qualitative questions aim at getting an in-depth, individualized, and contextually sensitive understanding of reading for each child interviewed. ... the purpose is to hear children talk about reading in their own words; find out about their reading behaviors, attitudes, and experiences; and get them to tell stories that illuminate what reading means to them. (Patton 2017, p. 14)

A number of participants commented on the value of qualitative approaches for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people. For example:

... 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. (Tasmanian Government, sub. 100, p. 4)

Evaluators should ensure that research objectives, standpoints and methodologies are driven and guided by Aboriginal and Torres Strait Islander communities. Aboriginal and Torres Strait Islander people take a holistic view on wellbeing, which can be poorly understood through discrete quantitative measures. Thus, taking a qualitative, storytelling and story sharing approach is important, as it captures Aboriginal and Torres Strait Islander 'ways of knowing'. (Riverina Murray Regional Alliance, sub DR149, p. 12)

A literature review on culturally competent evaluation by Chouinard and Cousins (2007), noted that a number of studies reported using qualitative methods (such as focus groups and interviews) to integrate Indigenous way of knowing into evaluations. These methods were used as a means of engaging participants in a reflective dialogue about the issues that mattered to them and their communities.

Muir and Dean citing Chouinard and Cousins 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 Liamputpong commented that, '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 4.9 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.

<i>Method</i>	<i>Description</i>	<i>Pros and cons</i>
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 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 and Wilson (2012, pp. 330–349).

Narratives and stories are part of sharing Aboriginal and Torres Strait Islander ways of knowing (box 4.10). Relationality is inherent in storytelling and can help participants to feel

respected as equal partners. Similarly, careful listening is inherent in yarning which can remove the power imbalance in a researcher-subject relationship (Drawson, Toombs and Mushquash 2017, p. 14).

Box 4.10 **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 is known by different names in different Aboriginal cultures, but it translates as 'contemplation'. It requires researchers to continually be reflective of their relationships with others and as an approach is empowering as it 'enables working with indigenous people and allowing their voices to be heard' (West et al. 2012, p. 1585). 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).

Yarning is mentioned as part of the evaluation methods for several of the Australian Government evaluations supplied to the Commission (appendix B provides detailed results).

Sources: Atkinson (2002, pp. 15, 20); Geia, Hayes and Usher (2013, p. 16); Leeson, Smith and Rynne (2016, p. 8), West (2012, p. 1585).

As with quantitative methods, qualitative data collection also needs to be representative of the population affected by the policy or program, and reflect different perspectives, to create

an overall picture of what is working or not working and why. While personal stories have long been an important part of evaluations, as BetterEvaluation said ‘an individual narrative is like a fragment of data that provides a perspective at one point in time from a particular point of view’ (BetterEvaluation 2010).

For most evaluations, a mix of quantitative and qualitative methods and sources of data will be needed to provide a complete picture of the outcomes of a policy or program.

Triangulation and the quality of evidence

Triangulation across sources of data can make each data source more valuable and strengthen 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 deliberatively 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 (table 4.3).

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).

Where a policy is relatively simple, experimental, quasi-experimental or statistical approaches can answer questions about whether and to what extent a policy worked. Where there is little understanding about how the policy will deliver impact, participatory and adaptive approaches can be used to generate information to provide real time feedback on the policy. *The Magenta Book 2020, Supplementary Guide: Handling Complexity in Policy Evaluation* provides a comprehensive list of evaluation questions and suitable approaches and methods for answering evaluation questions (HM Treasury 2020a, pp. 44–45).

Decision about what the evaluation questions should be, and what approaches and methods can be used to answer the evaluation questions, should be made with those expected to be affected by the policy or program and the potential users of evaluation (to maximise the usefulness and use of evaluation findings). A number of participants pointed to the importance of engaging with Aboriginal and Torres Strait Islander people on evaluation approaches and methods (and developing a theory of change). For example, The Interplay Project said:

In order to anchor Aboriginal and Torres Strait Islander peoples' perspectives, knowledges and cultures in our approach to Indigenous evaluations, we begin with capturing these worldviews and applying these into a Theory of Change, which we refer to as a Story of Change. This process can build from existing program logics or theories of change, which are typically developed by and from the perspective of program funding bodies and program owners rather than from the grassroots community perspective. Interplay's Story of Change provides a practical approach to centre and value Indigenous perspectives into articulating what communities see are steps that lead to effective change and how such change and outcomes should best be measured. (sub DR132, p. 3)

Table 4.3 What type of evaluation? It will depend on the purpose and the questions you want answered

Evaluation type	Questions	Purpose
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 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 assess 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 and approaches 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, for example, often also want to know what works, when and in what circumstances.

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, it 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 six-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).

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 and ensure adequate coverage of Aboriginal and Torres Strait Islander perspectives (box 4.13).

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)

Independent Members of the National Indigenous Australians Agency Indigenous Evaluation Committee commenting on the importance of flexibility in choice of methods said:

... evaluation is seen as being about both theory and data and therefore many methods hold the potential to demonstrate causality. (sub DR115, p. 3)

While Indigenous methods have largely been aligned with qualitative research, quantitative methods are increasingly being used for Indigenous research if, as Walter and Suina said, 'the data used reflect what is important to Indigenous peoples and reflect the ownership of data by Indigenous peoples' (2019, p. 233). Katz et al. also noted that:

In terms of the theoretical and methodological basis for Indigenous research, the literature tends towards constructivist or transformative approaches. These approaches to research and evaluation favour active participation or control by community members of the research methods

and dissemination. However, quantitative and economic evaluation methods are also increasingly being included as part of Indigenous research and evaluations when they meet the standards of Indigenous research; in particular when the community has a say in, and ideally control of, the research questions, methods and dissemination processes. (2016, p. 37)

The time and resources available for evaluation are also important considerations when choosing evaluation approaches and methods. RCTs and quasi-experiments, for example, can be (but do not need to be) expensive to run.

Are some types and approaches more or less suitable?

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 extensively 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 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).

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 RCTs²⁹ 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.

²⁹ BetterEvaluation (sub. DR144, pp. 11–12) also points out some of the potential risks of using experimental designs.

Figure 4.1 When to use quantitative impact evaluation

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

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

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 (2014).

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 4.4 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, adaptive learning and an agile approach to implementing policy

-
- allow Aboriginal and Torres Strait Islander knowledges, perspectives and world views to be valued and incorporated into the design and delivery of evaluations (including local ways of knowing, being and doing)
 - 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 4.10)
 - can help government agencies build relationships and trust with Aboriginal and Torres Strait Islander people and communities as they work together to find solutions (this aligns with the concept of reciprocity)
 - can strengthen the evaluation capabilities of Aboriginal and Torres Strait Islander people and communities (which in turn empowers and facilitates Aboriginal and Torres Strait Islander- and community-led solutions).

Community First Development, for example, said they co-design, monitor and evaluate all projects using a participatory approach:

... communities tell us what success looks like in the short term (projects), and the long term (dream). We then work closely with communities to track progress towards these goals. ... One of the driving purposes for selecting this approach was that it aligned with our ethical principles for communities to hold decision-making powers on whether projects should proceed and how they are designed, delivered and evaluated. It is inclusive and creates the space to build trust where honest answers are provided to inform project design and evaluation. A participatory approach can also accommodate First Nations evaluation methods where this is preferred. . (sub DR134, pp. 7–8)

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) (Westhorp 2014, p. 7).

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)

Co-design and participatory monitoring and evaluation practices should be foremost in the thinking and doing when working with First Nation peoples. Evaluation should be flexible and adaptive to meet needs within communities. (sub. DR174, p. 2)

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)

Developmental evaluation was designed for complex, innovative policies and programs (and relies heavily on program participants' responses to shape learning and improvement). 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)

Participatory and adaptive or 'learning by doing' approaches may be particularly useful in Aboriginal and Torres Strait Islander communities, where both local context and supporting capability building are important. These approaches allow Aboriginal and Torres Strait Islander knowledges and worldviews to be incorporated into evaluation design and the use of culturally sensitive methods of engagement. However, they will be less feasible when the evaluation findings need to be rigorously objective and evaluators need to be independent (HM Treasury 2020a, p. 49).

5 Lifting evaluation quality: Early planning and engagement

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.
- High-quality evaluations rely on early planning (setting out a clear program logic and resourcing needs), effective engagement and ethical conduct, rigorous and fit-for purpose evaluation methods, good data, quality assurance and good governance arrangements, and transparent evaluation methodologies, processes and findings.
- Evaluation is most effective when it is integrated into the design of a policy or program. Evaluation could be embedded into policy design by requiring Australian Government agencies to prepare an Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plan (IAEP) for new policies or programs proposed for government endorsement. The IAEP would address key questions on the planned engagement, evaluation approach, and resources needed for the proposed policy or program.
- Engaging effectively with Aboriginal and Torres Strait Islander people, including as partners in shared decision making across the evaluation process, will mean evaluation results capture what is valued by, the expertise, and lived experiences of Aboriginal and Torres Strait Islander people. This will improve the quality of evaluations and make them more useful.
 - Aboriginal and Torres Strait Islander leadership, for example through co-design, can play an important role in effective engagement for evaluations. However, co-design may not be appropriate in all circumstances. Co-design can be time and resource intensive for all participants, and requires government agencies to ensure that Aboriginal and Torres Strait Islander people have the resources and authority to decide what gets evaluated and how.
- All evaluations should be conducted ethically. This requires government agencies to carefully consider how ethical risks are assessed, as well as documenting how Aboriginal and Torres Strait Islander people influence this assessment, including how decisions are made on whether a formal review by an ethics committee is required.

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 poorly, evaluations can produce misleading results, leading to wasteful or potentially harmful policies and programs being continued. As the Organisation for Economic Cooperation and Development (OECD) 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. (2020c, p. 63)

An important reason for the 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 briefly at what it means to evaluate for high-quality results (section 5.1), before focusing on:

- early planning in practice, and how it could be embedded into policy or program development (section 5.2)
- engaging with Aboriginal and Torres Strait Islander people to ensure their perspectives, priorities and knowledges are centred in policy and program design, and evaluation practice (section 5.3).

5.1 What makes for a high-quality evaluation?

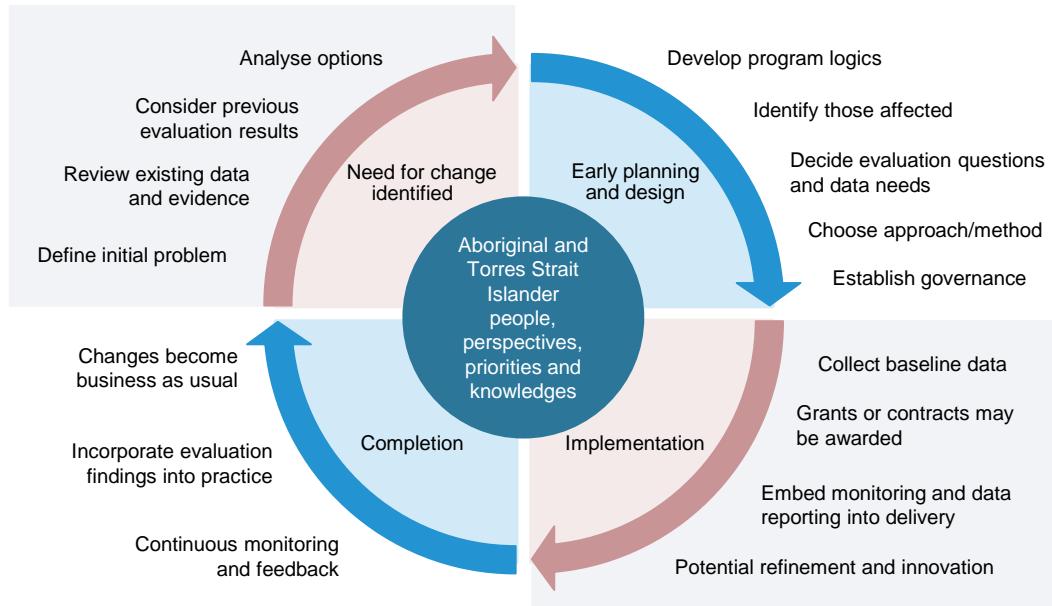
Evaluation is most effective when it is integrated into each stage of policy and program development, from setting policy objectives and collecting baseline data, through to using evaluation findings to inform future policy and program design (figure 5.1).

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)

Figure 5.1 Evaluation and engagement should be considered at every stage of the policy and program cycle



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 as highlighted in the *Magenta Book*, ‘small changes in … [policy or program] … 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).

Engaging with Aboriginal and Torres Strait Islander people and communities in policy design and at the various stages of evaluation ensures that ‘quality’ is not valued narrowly through an assessment of the process or design followed, but also captures the perspectives of those whose lives the policies and programs are designed to improve. This highlights the importance of the overarching principle of the Strategy, and involves a fundamental shift in how evaluations are thought about across the entire policy cycle. It involves:

- early planning and design — planning evaluations in partnership with Aboriginal and Torres Strait Islander people, including deciding on evaluation questions, approaches and methods (and other key stages of the evaluation process)

-
- implementation — ensuring evaluators have the necessary technical and cultural skills and experience in conducting evaluation in partnership with Aboriginal and Torres Strait Islander people
 - completion — translating and communicating evaluation findings with Aboriginal and Torres Strait Islander people and communities
 - identifying a need for change — keeping enduring partnerships with Aboriginal and Torres Strait Islander people to ensure evaluation findings are reflected in government priorities and policy changes.

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 (box 5.1 provides an example of a program logic for an anti-smoking initiative for young people).

Early planning and engagement are key

Early planning is key to quality evaluations and is critical to ensuring the evaluation delivers useful findings. It is important for ensuring evaluations identify the objectives and outcomes of the policy or program, and clarify what the evaluation questions should be (table 5.1 sets out questions to consider at the planning stage of an evaluation). Early planning can also help build a culture of evaluation as evaluation becomes embedded in the policy cycle rather than being 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 the collection of baseline data and finding an appropriate comparison group). It is also one way of ensuring that the 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 those effects are collected. This includes considering whether outcomes or access to policies or programs are different for Aboriginal and Torres Strait Islander and non-Indigenous people.

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, evaluation approaches and

methods, and data collection. It means thought can be given early on how to reach and engage with particular communities or groups, how to engage in a culturally safe way, and ensure the necessary time and resources are factored into the evaluation for such engagement.

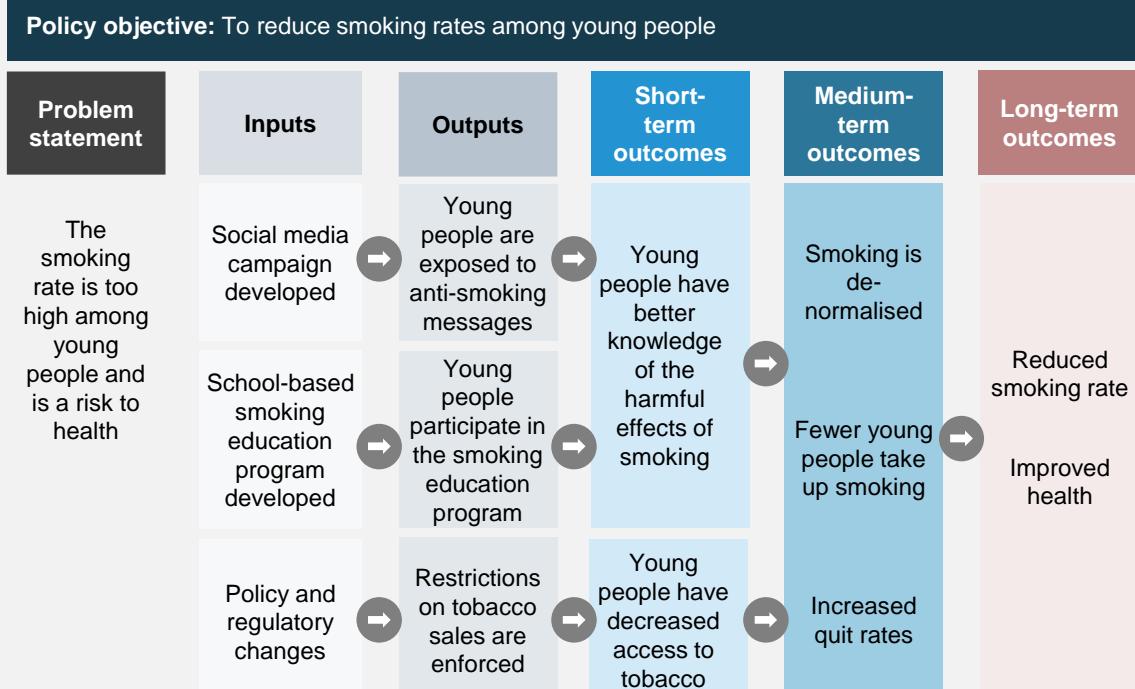
Early planning and engagement, including the ethical conduct of evaluations, are discussed in more detail in sections 5.2 and 5.3.

Box 5.1 Developing a policy or program logic model

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 and assessed during the evaluation (HM Treasury 2020c, pp. 24–26; OECD 2020c, p. 65).

Engagement with impacted populations in developing the logic model ensures outcomes and performance measures reflect what is valued by those whose lives the policy or program is intended to improve. This was highlighted by the Australian Evaluation Society:

Developing the theory of change ... 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)



Source: figure adapted from Centre for Epidemiology and Evidence (2017, p. 11).

Table 5.1 Questions to consider at the planning stage of an evaluation

The main steps in evaluation	Questions to consider
Define the policy objectives and outcomes to identify what to evaluate	What is the program logic or theory of change? Is the program logic clear on how the program outcomes are expected to be achieved and how these are linked to higher-level outcomes? Do the intended outcomes align to those identified by 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.) What is the likely impact on Aboriginal and Torres Strait Islander people or communities?
Identify the evaluation questions	What do policy makers and those affected by the policy or program need to know? (Have evaluation questions been decided on with Aboriginal and Torres Strait Islander people and other main users of the evaluation results?)
Select an evaluation approach and method(s)	What approaches and methods will answer each of the evaluation questions (will they require a formative, process, impact evaluation, or some combination)? Have the approaches and methods been decided on with those affected by the policy or program? Have contextual factors that are likely to influence outcomes been identified?
Identify data requirements (quality evaluations require good data)	What data are required? What is already being collected/available? What processes need to be put in place to allow the data to be collected? (If undertaking an impact evaluation, how will baselines be collected and when will impact be measured?) Are the data or the collection methods used responsive to the needs, rights and cultural safety of Aboriginal and Torres Strait Islander people?
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 Aboriginal and Torres Strait Islander people 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? Will the evaluation best be co-designed or Indigenous-led?
Use and disseminate the evaluation findings	What will the findings be used for? 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).

Technical rigour

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 types, 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 used should also be culturally responsive to reflect what is valued, and how evaluation is approached, by Aboriginal and Torres Strait Islander people. As Social Ventures Australia noted:

... 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. 4)

For qualitative evaluations, technical rigour relates to how evidence is gathered and analysed, used to draw findings, and 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 (discussed below).

Evaluation rigour can be enhanced by combining evaluation approaches and methods (such as experimental or quasi-experimental approaches with theory-based approaches) and using both quantitative and qualitative data (chapter 4).

Recognising and addressing bias

Evaluations can have biases that can compromise quality. 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)

Evaluation results, for example, can be affected by: 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.

For evaluations to be credible, there needs to be a degree of objectivity. This can be achieved through establishing independence in the process of conducting evaluations. As Patton said:

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)

For policy and program evaluation, several aspects 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 the 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 2020c, p. 67; UNEG 2016b, 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. And it is expected that the involvement of policy and program staff will enhance 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).

Validity of evaluation results

Quality evaluation results can be used to make better decisions on policies and programs. Validity of evaluation results helps to ensure potential biases are eliminated, and causal relationships are established to understand what policies and programs work. This includes establishing that '(1) the cause precedes the effect, (2) the cause is related to the effect, and (3) other plausible alternative explanations of the effect can be excluded' (Farrington 2003, p. 51).

Methodological validity can be assessed in several ways but two key criteria include: internal validity (referring to the robustness of the question being asked in the evaluation, for example, based on the quality of the control group); and external validity (referring to whether the estimate would hold up under non-experimental conditions when the intervention is scaled up or rolled out).

While understanding what works is important, ‘realist’ approaches to evaluation are about a broader measure of validity that is required to understand what works, for whom and in what context. This can be particularly important in framing evaluations that centre Aboriginal and Torres Strait Islander people and communities. La France et al. noted that:

Genuine understandings are grounded in place, setting, and community. Experiential justifications of validity, those grounded in the lived experiences of participants, are central to an indigenous framework. … knowledge is derived from individual and communal experience in daily life, in keen observations of the environment, and in interpretive messages received from spirits in ceremonies, visions, and dreams … Local language may be required to capture the meaning most closely related to the English-language expression of validity. (2012, p. 70)

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.

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. Valid sampling designs and data analysis techniques are important for technical rigour. This includes ensuring appropriate sampling strategies and sample sizes.

Chapter 9 looks at data issues in more detail.

Quality assurance and good governance

Quality assurance is also important for quality evaluations (HM Treasury 2020c, p. 75; OECD 2020c, p. 82). Ideally, quality assurance should occur throughout an evaluation so that evaluation design, planning and delivery are conducted systematically, and professional and analytical standards are adhered to.

Quality assurance can be conducted through steering committees, such as evaluation governance committees (box 5.2). 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.

Box 5.2 Evaluation governance committees

The main responsibilities of evaluation governance committees include:

- ensuring that the needs of decision makers and those affected by the policy or program 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.

Source: Department of Finance (1994a, pp. 37–38).

Quality assurance can also take place after an evaluation is completed by peer review or an independent assessor — adding another layer of accountability.

Quality assurance requires good governance arrangements. An essential part of these arrangements in the context of the Strategy is ensuring that partnerships with Aboriginal and Torres Strait Islander people are developed so Aboriginal and Torres Strait Islander people are at the table when determining what gets evaluated, how, when and by whom.

The National Indigenous Australians Agency's (NIAA) Indigenous Evaluation Committee provides Aboriginal and Torres Strait Islander oversight of evaluation activities within the Agency. The Committee is currently made up of two Aboriginal members, and two non-Aboriginal or Torres Strait Islander members 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).

The OCHRE (Opportunity, Choice, Healing, Responsibility and Empowerment) plan 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 first stage of 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 (NSW Ombudsman 2019, p. 40).

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 evaluation results, and keeps governments and decision makers accountable for their responses to evaluations (chapter 7).

Transparency can also help to build trust and instil confidence with key stakeholders, which is important if buy-in is needed to implement changes following an evaluation (UNEG 2016b, 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).

Time and resources

Quality evaluations require time and resources. Agencies should determine early, as part of their evaluation 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)

The cost of an evaluation can vary 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 linguistic 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.

A number of participants argued that evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people are often not adequately funded.³⁰ 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)

Engaging effectively with Aboriginal and Torres Strait Islander people and more participatory forms of evaluation can require additional time and resources, compared to other evaluations. For example, commenting on an 18-month long evaluation commissioned by LawRight of health justice partnerships, Fiona Allison and LawRight said:

Consultation and engagement take time to do well. Reasons for this include, in this instance, the complexity of the needs of and issues impacting on clients, meaning that more work is involved in ensuring that their input is gathered. Time has also allowed for relationships between those

³⁰ For example: Children's Ground (sub. 73, p. 8); the Western Australian Government (sub. 74, p. 2); and National Aboriginal Community Controlled Health Organisation (sub. 95, p. 5).

involved in the [health justice partnerships] and in the evaluation to grow, also likely to [have increased] the quality of consultation and engagement. (sub. 18, p. 3)

The Western Australian Government said that time and cost of establishing relationships 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 UNSW Social Policy Research Centre also spoke about the budgetary implications of long lead times to connect and consult with communities.

It is especially important to consider the budgeting and resourcing needs of Aboriginal and Torres Strait Islander evaluations because these have unique considerations. In our experience, evaluations with Aboriginal and Torres Strait Islander communities require a long lead time in order to connect and consult with communities about their interest in the evaluation and their priorities for research, and this does have implications on resourcing and budget. (sub. DR127, p. 4)

5.2 Early planning in practice

While all stages of an evaluation — from planning to completion and use — are important for lifting the quality of evaluation results, getting things right from the start is critical.

Yet, while over half of the agencies that responded to the Commission's information request said that evaluation plans were 'always' or 'very often' required when developing new policies and programs, about one-quarter said evaluation plans were 'never' or 'rarely' prepared in developing new policies and programs (appendix B).

The absence of good data, including baselines, was also a commonly cited reason why evaluations were unable to answer evaluation questions (chapters 3 and 9). This underlines the importance of thinking clearly through what the end goal of the policy or program is, and how its success will be measured (using a policy or program logic model, box 5.1) from the very beginning. 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)

Evaluation planning can also ensure the differential impacts of mainstream policies and programs on Aboriginal and Torres Strait Islander people (and other minority groups) are considered in the policy and evaluation design, and that the data required to evaluate these effects are collected.

Responses to the Commission’s information request showed that about two-thirds of evaluations of mainstream policies and programs did not consider their effects on Aboriginal and Torres Strait Islander people. A lack of data was also noted as a reason why an analysis of impacts for Aboriginal and Torres Strait Islander people was not undertaken (or was limited) in a number of the mainstream policy and program evaluation reports provided to the Commission. An example of an evaluation that was not able to answer questions about impact because of a lack of data is provided in box 5.3.

While evaluation questions should be guided by early engagement with Aboriginal and Torres Strait Islander people (section 5.3), evaluations of mainstream policies and programs should be, at the policy development and early evaluation planning stage, asking questions such as:

- is the policy or program likely to work differently for Aboriginal and Torres Strait Islander people, or differently for groups of Aboriginal and Torres Strait Islander people (or other groups)? And, why might it be working differently for different groups?
- are Aboriginal and Torres Strait Islander people likely to access the policy or program at the same rate as other populations?
- are there barriers to Aboriginal and Torres Strait Islander people accessing the policy or program? Will the policy or program be delivered in a culturally safe way?

Embedding early evaluation planning in policy development

Policy development rarely fits into a neat policy cycle. Changes to policies or programs can occur at any stage of their lifecycle, and critical steps in policy development can be missed because of time and/or budgetary constraints. One way to embed evaluation planning into policy development is to set a mandatory requirement for early planning and engagement to occur before any new policy or program can be considered for government endorsement.

Setting such a ‘gatekeeping’ requirement is a fairly blunt instrument, but it ensures compliance, and elevates the importance of early planning in the decision-making process — two important components that could assist in creating cultural change. On the downside, if the requirement is treated just as a compliance exercise, the benefits will be reduced.

New Policy Proposals are currently mandated to include an assessment of their impacts on the budget, on regional Australia, and of the regulatory costs they impose on the community. All decisions made by the Australian Government, including non-Cabinet decisions made by Ministers or statutory authorities (where the proposal is not a minor or machinery change), are required to assess their regulatory impact on businesses, community organisations or individuals through a Regulation Impact Statement (DPMC 2020a, p. 1).

Box 5.3 **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).

It is mandatory for a Regulation Impact Statement to be prepared as part of all Cabinet submissions. Cabinet submissions may also require the completion of Regional Australia Impact Statements if the proposal is expected to have a 'significant positive or negative impact on regional Australia' (DITRDC 2020, p. 3). According to guidance by the Department of Infrastructure, Transport, Regional Development and Communities, in assessing a proposal's significance:

Agencies may wish to consider proposed expenditure, the size of the population/geography impacted, the [Average Staffing Level] sought, impact on industries and economies, impact on social outcomes, or the subjective perspective of regional Australians ... (DITRDC 2020, p. 3)

A number of jurisdictions also have mandatory evaluation requirements in place. For example:

- all Western Australian Government 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 future funding decisions. 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)
- all Canadian Government federal 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).

As discussed in chapter 2, in the late 1980s and early 1990s, the Australian Government had formal evaluation requirements 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 were to include a statement of proposed future evaluation.

The Independent Review of the Australian Public Service did not explicitly support bringing back a system of mandated evaluation planning and activity, but rather proposed 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).

There was strong support from participants for New Policy Proposals³¹ 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 turn out to have a significant impact on Aboriginal and Torres Strait Islander people.

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

³¹ It is important to note that evaluation plans in New Policy Proposals only deal with the flow of new policies and programs, not the large stock of existing policies and programs (chapter 6 provides more detail).

-
- 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)

Early impact assessments of policies and programs on Aboriginal and Torres Strait Islander people are already undertaken in some states and territories in the health space, through the Aboriginal Health Impact Statement. In New South Wales, South Australia and Western Australia, the statement is mandatory for health agencies when developing new policies, programs or strategies, and must accompany the proposal as it is sent for endorsement (Department for Health and Ageing, Government of South Australia 2014; Department of Health, Government of Western Australia 2017; Ministry of Health, NSW Government 2017).

The requirements of the Aboriginal Health Impact Statement differ slightly in each jurisdiction, but in essence ask agencies to assess: the health and policy context for Aboriginal people associated with the proposal; the potential impact of the policy, program or strategy on Aboriginal people; and any engagement undertaken with Aboriginal people. According to NSW Health:

The Aboriginal Health Impact Statement is used by NSW Health organisations to ‘incorporate the health needs and interests of Aboriginal people in the development of new and revised health policies, programs and strategies’

Applying an ‘Aboriginal health lens’ to programs at an early stage aims to improve the health and wellbeing of Aboriginal people by ensuring the health needs of Aboriginal people are carefully considered. It also assists to identify opportunities to improve access to healthcare, achieve equity, develop effective partnerships and conduct meaningful consultations. (NSW Government nd)

Anecdotal evidence from stakeholders, and backed up by a mid-term evaluation of the NSW *Aboriginal Health Plan 2013–2023*, suggests adoption of Aboriginal Health Impact Statements varies across the health network (Centre for Epidemiology and Evidence and Centre for Aboriginal Health 2019, pp. 75, 83).

An Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plan

Because early planning and engagement is so important for the quality of evaluations, the Commission is proposing an ‘Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plan (IAEP)’ be required for all new policy or program proposals developed by Australian Government agencies for Cabinet or Government (in the case of decisions by Ministers and/or statutory authorities) endorsement (Strategy, action 3).

The IAEP would establish a standard for the level of planning and engagement appropriate for the proposal being prepared for government endorsement, based on the expected level of impact on Aboriginal and Torres Strait Islander people. It would include details on:

- the Aboriginal and Torres Strait Islander people, organisations and communities that would be affected by the proposed initiative
- a proposed engagement plan (including timeframes and cultural safety considerations)
- a proposed evaluation approach (including the scale of evaluation required)
- the data required to assess the policy's impact and how they would be collected
- an estimated evaluation budget (Strategy, action 3).

For policies where the impact on Aboriginal and Torres Strait Islander people is expected to be high, the expectation is the proposed engagement and evaluation plans, and evaluation approaches will reflect this. Chapter 6 discusses how the IAEP informs how the policy or program proposal would be evaluated under the Strategy.

The IAEP would codify good practice and be a check for decision makers that evaluation planning and engagement have informed the development of a new policy or program they are being asked to endorse.

While the IAEP would not be a public document, not adhering to the intent of the IAEP would be highlighted (through community feedback and assessment by the proposed Office of Indigenous Policy Evaluation) when key initiatives were not included in published Departmental Three-Year Forward Work Plans (chapter 6). Similarly, published evaluation reports that are not able to demonstrate how a clear program logic and engagement helped inform decisions about what was evaluated, how, and when, would highlight a failure of the agency to properly consider key aspects of early planning and engagement. Transparency, and the monitoring of agencies performance against the Strategy (chapter 10) will provide the incentive for agencies to elevate the role of early planning through the Aboriginal and Torres Strait Islander IAEP.

There may be some cases where all elements of the IAEP cannot be developed in time to be considered at the decision-making stage of a policy or program cycle. For example, in the case of a policy response to an unexpected event, such as a health pandemic. In such cases, the requirement for an IAEP could be replaced with a commitment to undertake a formative evaluation to ensure that key stages of early planning are undertaken at the earliest opportunity.

5.3 Effectively engaging with Aboriginal and Torres Strait Islander people

Quality evaluation relies on agencies engaging effectively with those affected by a policy or program. Many of the concerns raised 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' (box 5.4) — and the lack of regard for Aboriginal and Torres Strait Islander knowledges (box 5.5).

Box 5.4 Participants highlighted the importance of engagement for quality evaluation

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)

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)

Beyond Blue:

Working in partnership with Aboriginal and Torres Strait Islander people on evaluation activities of government funded programs, both Indigenous specific and mainstream programs, will ensure current and future Australian governments can support informed solutions that work and make progress in closing the gap. (brief comments, 2)

The National Mental Health Commission 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)

Box 5.5 Indigenous knowledges and worldviews

Indigenous or traditional knowledges refers to the complex bodies and systems of knowledge, know-how and practices developed and maintained by Indigenous peoples. These knowledge systems are based on values, beliefs, rituals and community laws and practices. They include concepts and methods for land and ecosystem management.

Indigenous knowledges tend to be collectively owned and are often transmitted orally from one generation to another. They can take the form of stories, songs, beliefs, customary laws and artwork, as well as more technical knowledges and skills. The transmission of traditional knowledges is a practical process of continuing and sharing knowledges, and preserving Indigenous cultures, languages and identities.

The value of Indigenous ways of knowing, being and doing as an input into policy design and evaluation is increasingly being recognised and acted upon. Some examples include using:

- Indigenous knowledges with artificial intelligence to solve complex environmental management problems (CSIRO, sub. DR131, p. 7)
- Indigenous customary fire management to reduce carbon emissions from wildfires in northern Australia (Fisher and Altman 2020).

As the United Nations Economic and Social Council explained:

There is increasing recognition of the unique value of [Indigenous] peoples' traditional knowledge and its potential to contribute to addressing some of the most significant challenges faced today, including achieving sustainable development, mitigating climate change, managing conservation areas and helping in the development of new technology and medicine, such as pharmaceuticals. (2019, p. 2)

Others, however, suggest that Indigenous knowledges are not always considered to be of equal value to Western knowledges and dominant worldviews. Smith, for example, said:

The struggle for the validity of indigenous knowledges may no longer be over the that [Indigenous] peoples have ways of viewing the world which are unique, but over proving the authenticity of, and control over, our own forms of knowledge. (1999, p. 104)

And some participants suggested that Indigenous knowledges are not always heard:

- 'Aboriginal and Torres Strait Islander people can say things and those listening cannot actually 'hear' what is being said' (Maggie Walter, sub. 112, p. 2)
- The Kimberley Aboriginal Law and Cultural Centre and Cultural Development Network said that a key threshold issue for the Strategy is 'the recognition that Indigenous Australians have an existing knowledge that is ignored at best, and in most cases, actively worked against' (sub. DR129, p. 2).

The Institute for Human Security and Social Change noted the potential challenges for evaluation when there are different values and worldviews.

It is the experience of the Institute that the worldviews and values which are encountered in consultation with Aboriginal and Torres Strait Islander people are often different to the values and worldviews of government and other white Australian stakeholders, creating inherent tensions and challenges for evaluation framing and implementation. (sub. DR119, p. 2)

Engaging with Aboriginal and Torres Strait Islander people and communities in policy and program design and evaluation is key to ensuring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred in evaluation (table 5.2). This

helps to deliver evaluations (and policies) that are meaningful to, and valued by, Aboriginal and Torres Strait Islander people.

Table 5.2 Engaging at each stage of an evaluation

<i>Engaging with Aboriginal and Torres Strait Islander people is important when ...</i>	
... defining policy objectives and outcomes and deciding what to evaluate	to ... <ul style="list-style-type: none">ensure the program logic or theory of change reflects the outcomes that are valued by Aboriginal and Torres Strait Islander people.
... planning an evaluation and deciding on evaluation governance arrangements	to ... <ul style="list-style-type: none">ensure evaluation objectives and questions, methodologies, and ongoing engagement arrangements reflect the expectations and needs of Aboriginal and Torres Strait Islander peopleallow adequate time and resources for data collection, ethical approvals, and engagement during the various stages of the evaluationselect a team for the evaluation with the required cultural and technical skills.
... collecting data and analysing the evidence	to ... <ul style="list-style-type: none">embed effective data collection and monitoring arrangements (including within organisations that deliver services for government)enable timely feedback during implementation.
... interpreting, reporting and using the results	to ... <ul style="list-style-type: none">validate or refine evaluation findings with evaluation participants in light of evidenceensure that findings about Aboriginal and Torres Strait Islander people are presented accurately and respectfully, and are shared with relevant communitiestranslate evaluation findings into policy and program improvements that will improve the lives of Aboriginal and Torres Strait Islander people.

The term ‘engagement’ is used (and misused) to refer to many types of interactions, but can be defined as an interaction between groups of people working towards shared goals. The literature refers to engagement as not being a single process or set of activities, but rather an ongoing process or conversation that builds trust and relationships (Hunt 2013, p. 5). In the context of engaging with Aboriginal and Torres Strait Islander people, 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)

This definition of effective engagement mirrors feedback from participants of the importance of engaging in a way that builds longer-term relationships based on trust and respect, and that allows Aboriginal and Torres Strait Islander people to be equal partners in the evaluation process. The importance of recognising the heterogeneity of Aboriginal and Torres Strait

Islander people, perspectives, priorities and knowledges, and ensuring that many voices are heard during the engagement process was also emphasised.

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

Lowitja Institute Evaluation Framework for Aboriginal and Torres Strait Islander Health	NHMRC Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities	AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research	Kaupapa Māori evaluation principles (New Zealand)^a
<ul style="list-style-type: none">• 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	<ul style="list-style-type: none">• Spirit and integrity• Cultural continuity• Equity• Reciprocity• Respect• Responsibility	<ul style="list-style-type: none">• Indigenous self-determination• Indigenous leadership• Impact and value• Sustainability and accountability	<ul style="list-style-type: none">• 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• Do not trample the mana (power, dignity and respect) of the people• Be humble in your approach

^a This table presents one interpretation of Kaupapa Māori evaluation principles, based on Cram (2009).

Sources: AIATSIS (2020b); Cram (2009); Kelaher et al. (2018a); NHMRC (2018a).

Engagement helps produce evidence of value

Because evaluation is an exercise in assessing the value or merit of a policy or program (chapter 1), the metrics used to measure 'value' in an evaluation should reflect the lived experiences of those directly affected by the policy or program.

Engagement processes can help to articulate outcomes that are valued by Aboriginal and Torres Strait Islander people. As Walker et al. said:

... evaluations which are attempting to assess the extent to which programs, policies and practices contribute to strengthening community (in both Indigenous and broader societal contexts) need to take into account the extent to which Indigenous goals, terms of reference, equitable processes and outcomes are achieved. An evaluation of such scope requires particular process oriented research approaches, including models based around participatory action research or community education. It also requires moving beyond purely efficiency and effectiveness measures. (2002, p. 23)

Some participants suggested that evaluations often overlook the extent to which policies and programs represent value for Aboriginal and Torres Strait Islander people. The RMIT 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)

Others pointed out the importance of considering cultural value in policy and program design. 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.

As an example, kidney dialysis in a hospital (away from the communities of patients) has benefits, but the social costs of patients being separated from Country (including disruption to kinship and cultural ties, and not being able to fulfil cultural commitments) may outweigh these benefits. A mobile dialysis truck that visits remote communities allows patients to have dialysis on Country, and achieves both benefits (box 5.6).

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). Aboriginal and Torres Strait Islander evaluation methods and approaches are also helpful in understanding value from the perspectives of Aboriginal and Torres Strait Islander people (for example, Kelaher et al. (2018a, 2018b) and Williams (2001), and chapter 4).

Effective engagement that is able to capture evidence that is valued 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. In many cases building this expertise will require investment in internal capacity and governance (discussed below).

The importance of early engagement

Meaningful engagement is about people being able to shape policy or program objectives and outcomes, as well as the evaluation questions, approaches, methods, and data collection. And this requires engaging with those affected by a policy or program at the earliest point in the policy or program (and evaluation) cycle. 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).

Effective engagement requires adequate time, resourcing and cultural, linguistic and technical capability be built into the evaluation. Engagement should be a key plank in policy (and evaluation) planning so there is a budget for it. Trying to retrofit some form of engagement later into the policy process may be too late to collect data on meaningful outcome measures, and this will impact the value of the evaluation and/or result in more expensive evaluations.

Engagement will be most successful when those involved are clear about what they want to achieve and who is responsible for what. This requires appropriate governance arrangements. The NIAA highlighted the importance of Aboriginal and Torres Strait Islander people playing a role in deciding on governance structures to ensure they are culturally appropriate.

Ultimately, Indigenous communities and organisations should play the key role in determining what governance structures and composition forms are suitable to ensure these are culturally appropriate. This needs to involve active promotion of participation by Indigenous academics and Indigenous-led institutions, organisations or companies to participate in evaluation, and local community members to participate in data collection. This increases ownership of evaluation, and improves the potential for culturally responsive evaluation practice. (sub. DR161, p. 2)

As discussed earlier, some government agencies (including the NIAA) already have governance arrangements in place to ensure they are responsive to the needs of Aboriginal and Torres Strait Islander people, including through Aboriginal and Torres Strait Islander representation on evaluation governance or steering committees. To meet Australian Government commitments on priority reforms under the National Agreement on Closing the Gap, other agencies will need to look at establishing formal partnership or shared

decision-making arrangements with communities or Aboriginal and Torres Strait Islander community controlled organisations.

Box 5.6 **An 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 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 Mobile 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 Mobile Dialysis Truck).

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

In practice, good governance arrangements will also help to prioritise agency evaluation activity to areas of greatest value, including the level of engagement required. Getting governance right ensures that engagement is framed in a culturally safe, and meaningful way

from the start. This is particularly important in mainstream settings where cultural capability may not be as well developed (or front of mind during decision-making processes).

Governance arrangements can also support a responsive approach to engagement. For example, an Aboriginal and Torres Strait Islander evaluation steering committee could decide that Aboriginal and Torres Strait Islander people should lead the development of outcome indicators and the selection criteria for contracting out the evaluation but that other parts of the evaluation require more limited engagement.

The continuum of engagement

Engagement involves costs for both governments and the people and communities involved.³² The level of engagement for any evaluation therefore needs to reflect the value or benefit which government and/or the community places on the engagement. A policy or program that is assessed as being a priority for improving the lives of Aboriginal and Torres Strait Islander people should consider implementing more involved forms of engagement, where the community has influence over decisions in the evaluation. Policies and programs assigned as a lower priority, on the other hand, may require less ongoing engagement, where the community opts out from being directly involved at every stage of the evaluation.

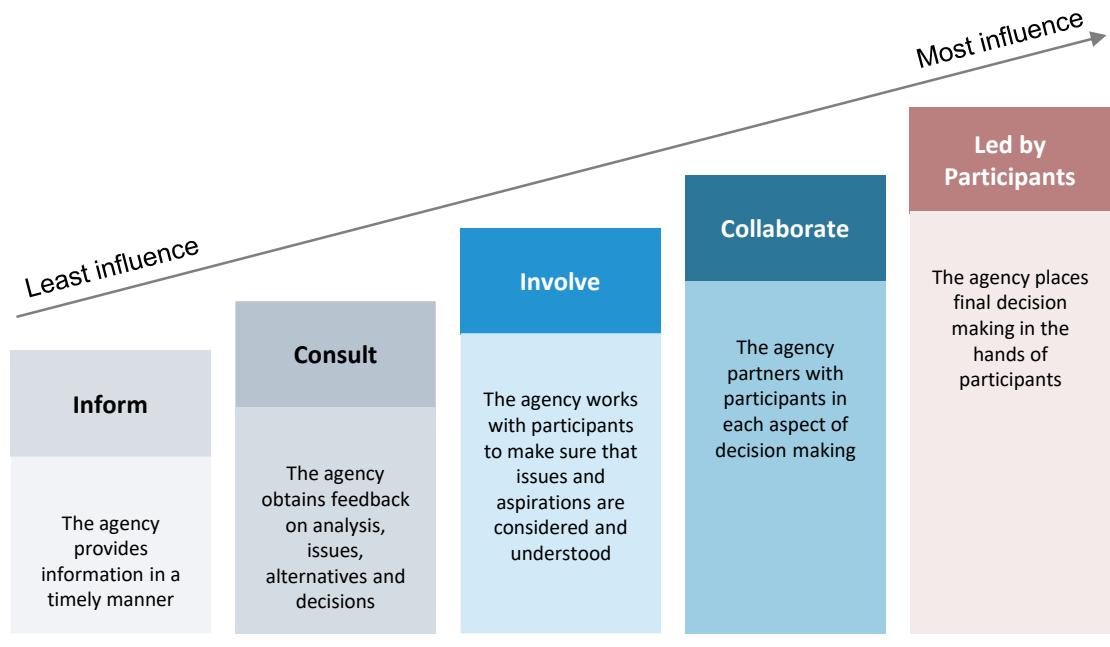
The literature refers to a continuum of engagement, which ranges from informing to participant-led (also known as empowerment evaluation, chapter 4) (figure 5.2). Engagement at the higher end of the continuum can be particularly helpful when evaluating complex problems and policy areas where working collaboratively with participants can help find policy solutions (Hunt 2013). And for many policies and programs affecting Aboriginal and Torres Strait Islander people, complexity and unknown solutions are factors in play. For example, 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)

While the continuum is a helpful stylised approach to engagement, in reality an agency could adopt a number of approaches along the continuum for any given evaluation. Agencies also need to make decisions about the breadth of engagement (that is, how widely or narrowly to engage).

³² A number of participants commented on the unrecognised burden often experienced by Aboriginal and Torres Strait Islander service providers who are required to participate in evaluations (for example, VACCA, sub 27, p. 2).

Figure 5.2 A continuum of engagement



Source: adapted from International Association for Public Participation (2018).

Co-design and community-led engagement

Co-design

Many participants pointed to the advantages of ‘co-design’ as an important approach to centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges in evaluations; and recognising that Aboriginal and Torres Strait Islander people should be involved in decisions that affect them, including the design, delivery and evaluation of policies and programs (box 5.7). 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).

Co-design reflects higher levels of influence along the engagement continuum — within the collaborate and participant-led end (figure 5.2) — although in practice the term co-design has been applied much more broadly.

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 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)

Co-design 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).

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

National Voice for our Children (SNAICC) explained that genuine co-design requires:

... a commitment for agencies to engage the perspectives of Aboriginal and Torres Strait Islander people from the very beginning of the evaluation spectrum to the very end. ... [for example] ... [a]gencies ... to ensure that Aboriginal and Torres Strait Islander communities are afforded an opportunity to co-design evaluation methodologies at the inception phase, rather than as a consultative component to inform research and evidence. This includes the opportunity to define the outcomes and measures that are important to Aboriginal and Torres Strait Islander people, rather than only responding to government defined goals and targets. (sub. DR163, p. 7)

Such partnering arrangements can strengthen evaluation capability and facilitate a sense of joint ownership. They can also strengthen trust and the cultural capability of non-Indigenous evaluators and those delivering policies or programs.

The National Agreement on Closing the Gap outlines some important elements of a strong partnership with Aboriginal and Torres Strait Islander people (sections 32 and 33) (JCOCTG 2020, pp. 6–7), which includes adequate funding ‘to support Aboriginal and Torres Strait Islander parties to be partners with governments in formal partnerships’ (JCOCTG 2020). Some participants highlighted the importance of establishing a formal agreement or terms of reference to overcome common issues with co-design. For example, the Alliance of First Nations Independent Education and Training Providers noted some common challenges in co-design including: agreeing to common goals before the process commences, which should drive decision making; determining the process that decisions will be reached by; and understanding and accepting the limitations of budgets allocated to a program. (sub. DR141, p. 2).

Maggie Walter noted that real co-design in practice is an equal share in decision making and equality of power, and for this to occur Aboriginal and Torres Strait Islander people need particular skills.

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. And this comes back to importance of rapid capacity-building and data expertise of Aboriginal and Torres Strait Islander people, which is still relatively rare in the public service. To legitimately participate in co-design, Aboriginal and Torres Strait Islander people need to have those skillsets in place. (sub. 112, p. 2)

Co-design can also improve transparency — participants are aware of (and engaged in deciding on) evaluation questions, 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.

An example of a co-designed evaluation 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,³³ which is co-designed with experts in Aboriginal and Torres Strait Islander health and with local

³³ 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 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 NSW 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. Some key 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 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).

Many of these key learnings of how community shapes evaluation design and outcomes through a genuine co-design process can also be seen in the example in box 5.8.

Co-design will not suit all evaluations. It requires considerable time and resources and for government agencies to be flexible. 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, pp. 24–25).

The extent to which agencies are able to implement co-design in their evaluations will depend on their capacity to adapt to policy or program changes as issues arise through the co-design process, as well as whether they have the authority to share decision making with Aboriginal and Torres Strait Islander people. As the Australian Institute for Primary Care and Ageing said:

Some co-design approaches are not appropriate on the given resources and may result in a superficial consultation that may not add value. ... Justification for a limited approach to engagement may relate to the scope of the project, the available budget, and the potential for community to influence direction and decision. That is — undertaking a co-design process without actually ceding any power is not really co-design; but ceding power can lead to outcomes that agencies might find difficult to accept. Being aware of how plausible and genuine a comprehensive codesign approach is prior to commissioning an evaluation is likely to be beneficial to all. (sub. DR139, p. 3)

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 children 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 an 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).

National Voice for our Children (SNAICC) noted:

Justification for a limited approach to engagement may relate to the scope of the project, the available budget, and the potential for community to influence direction and decision. That is – undertaking a co-design process without actually ceding any power is not really co-design; but ceding power can lead to outcomes that agencies might find difficult to accept. Being aware of how plausible and genuine a comprehensive codesign approach is prior to commissioning an evaluation is likely to be beneficial to all. (sub. DR163, p. 3)

The Alliance of First Nations Independent Education and Training Providers also said that it is important to acknowledge ‘that it is not always within the groups capacity to make some decisions’, for example where there are legislative requirements. (sub. DR141, p. 2).

Co-design will also require skilling up evaluators across agencies to ensure they are able to work collaboratively with participants. The types of changes that agencies would be required to make align with changes recommended previously by the Commission in the report on *Expenditure on Children in the Northern Territory* for agencies to adopt a more relational approach to contracting. The report recommended that staff and managers at a regional level have:

- skills in community engagement, including an understanding of the cultures and the communities they are engaging with
- capacity in terms of the time and resources to work with communities across the region

-
- autonomy to deliver impartial advice, including the authority to deal with issues that arise at a regional level. (PC 2020, p. 187)

The Western Australian Council of Social Service has developed a Co-design Toolkit, which includes questions to help decide when co-design might be the ‘right’ approach (WA Council of Social Service 2017, p. 7). It provides practical advice such as, co-design may be considered if a contract is due to expire within 18 months, but not if it expires in three months. Also, if you ‘already know what needs to be done and how to do it’, the value of co-design could be questioned.

Given the limited experience with co-design, more research and experience is needed to understand how co-design can better work 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’.

Aboriginal and Torres Strait Islander-led evaluations

Aboriginal and Torres Strait Islander-led evaluations are where Aboriginal and Torres Strait Islander people have final decision-making authority over the evaluation process. In practice, this can mean that:

- evaluators or leading members of the evaluation team are Aboriginal and Torres Strait Islander people
- evaluations and evaluation priorities are driven and guided by Aboriginal and Torres Strait Islander people and communities (NHMRC 2018a, p. 2).

It also means that Aboriginal and Torres Strait Islander standpoints, research methodologies and practices are valued, and used where appropriate.

As discussed in chapter 8, evaluations led by Indigenous people, are considered more likely to have the cultural capital to:

- facilitate respectful engagement and observe cultural protocols
- use Indigenous knowledges, methods and ways of working in an evaluation
- embed a cultural ‘reading’ or assessment of the cultural fit or appropriateness of data-collection methods and tools for Indigenous peoples
- facilitate what value and goodness look like through an Indigenous worldview
- lead sense making and analysis to ensure the richness, subtlety and nuance of meaning are not lost in translation and ensure the cultural validity of the evaluation conclusions (Wehipeihana 2019a, p. 372).

Many participants called for the Strategy to promote Aboriginal and Torres Strait Islander leadership in evaluation of policies and programs impacting Aboriginal and Torres Strait Islander people.

... authentic, considered, community led and driven evaluation is critical to improving the health outcomes for Aboriginal and Torres Strait Islander peoples. (Queensland Aboriginal and Islander Health Council, sub. DR173, p. 1)

If evaluations are to capture the evidence of progress, achievement and what works in Indigenous contexts and settings, then Indigenous peoples as experts in their lived experience, their community and country, need to be valued, affirmed and privileged. Indigenous peoples are vital to unlocking insights and generating understanding for use in policy and planning. Indigenous evaluation leadership and control is vital if Indigenous views and values are to hold sway. (Australian Evaluation Society, sub. DR174, p. 3)

Aboriginal and Torres Strait Islander leadership was also seen as a way to address cultural capability gaps. Tony Holmwood commented that:

Analytical or rational specialists tend to have a limited concept of cultural or social issues, so it is important aboriginal leaders and elders are involved in monitoring the outcomes. And this is also a matter of respect so programs are not imposed on the communities. ... Building partnership and trust, and associated evaluation techniques, needs to happen from the ground up with the focus on building purposeful communities and meaningful social connection, that will in turn grow community and leadership confidence and competence. Top-down evaluation techniques will fail if they are not well targeted to these objectives. (sub. DR175, pp. 1, 3)

Evaluations that are focused on 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)

Greater Aboriginal and Torres Strait Islander leadership in evaluation requires government agencies to think differently about what constitutes ‘expertise’ in evaluation, and to have the authority to empower communities to have greater agency over what gets evaluated and how. The National Centre of Indigenous Excellence argued that:

... the Strategy must create opportunities for First Nations' agency in problem-defining and solution-making. There must be a shift from a 'government led - knows best' (and in its worst form 'saviour') mentality to enabling and empowering Aboriginal and Torres Strait Islander people to create solutions for their own communities. (sub. DR158, p. 3)

Chapter 8 looks at ways to strengthen cultural capability by increasing Aboriginal and Torres Strait Islander leadership in evaluation of policies and programs that impact Aboriginal and Torres Strait Islander people. The Interplay Project Pty Ltd, for example, saw a need to 'invest in the Aboriginal and Torres Strait Islander evaluator workforce and create leadership opportunities that enable Aboriginal and Torres Strait Islander people to drive and influence evaluation practice' (sub. DR132, p. 4).

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.

Evaluations 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 2013c; HM Treasury 2020c, pp. 76–79; Posavac and Carey 1997, pp. 85–101).

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 Dhariwaa 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 Aboriginal and Torres Strait Islander researcher capability, and opportunities for recognition and employment of Aboriginal and Torres Strait Islander 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 mentees) 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 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, Dhariwaa Elders Group and UNSW (sub. 38, attachment).

Ethical practice in 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' (AIATSIS) *Code of Ethics for Aboriginal and Torres Strait Islander Research* (2020b) (and associated *A Guide to Applying the AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research* (2020a))
- National Health and Medical Research Council (NHMRC) guidelines on *Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities* (NHMRC 2018a).

Ethical principles in human research in Australia are also guided by the *National Statement on Ethical Conduct in Human Research* (National Statement) (NHMRC, ARC and UA 2018). 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. (2018b, p. 1)

To ensure ethical requirements are upheld, evaluations can be subject to formal ethics review. This 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)) (box 5.10). Human Research Ethics Committees (HRECs) 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.

According to the National Statement, all research with Aboriginal and Torres Strait Islander people is subject to ethics review by a HREC, and this must include assessment or advice from:

... people who have networks with Aboriginal and Torres Strait Islander Peoples and/or knowledge of research with Aboriginal and Torres Strait Islander Peoples; and people familiar with the culture and practices of the Aboriginal and Torres Strait Islander people with whom participation in the research will be discussed. (NHMRC, ARC and UA 2018, p. 78)

The updated *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research* (2020b), however notes that this requirement is likely to be reviewed and changed to be consistent with a more proportionate approach. The AIATSIS Code further:

... recognises that any assessment of risks and benefits of research must be considered in the context of the collective rights and interests of Indigenous peoples in relation to their lands, waters, cultures and histories. (AIATSIS 2020b, p. 7)

Box 5.10 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).

Many participants highlighted the importance of following a robust ethical process for all evaluations of policies and programs that impact Aboriginal and Torres Strait Islander people. However, balancing this was a desire to ensure that the costs of the ethics process does not unduly get in the way of evaluation activity (box 5.11). For example, the Australian Institute for Primary Care and Ageing noted:

Too often evaluators and researchers exclude vulnerable groups because they want to avoid restrictions and delays that can be associated with ethics approval processes. Requirements to

have explicit ethics approval to gather data from Aboriginal and Torres Strait Islander people can be misinterpreted to mean that they should be excluded if approval is not sought for their inclusion. This in effect can disempower Aboriginal and Torres Strait Islander people as their views cannot be heard or validated when they are part of a population. (sub. DR139, p. 3)

Box 5.11 Participants supported ethical evaluation practice, but some also highlighted 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)

Queensland Aboriginal and Islander Health Council:

... all evaluations commit to and are accountable to rigorous ethics review and show evidence that core values of the research guidelines are upheld. [the Queensland Aboriginal and Islander Health Council] notes that the Strategy has mentioned that not all evaluations will be subject to an ethics review. It is important that the Strategy adheres to the core values of research guidelines, completely, and every evaluation must follow the rigorous and appropriate ethics review process. This will ensure that research guidelines are upheld and that there is minimal risk to Aboriginal and Torres Strait Islander communities and individuals. (sub. DR173, p. 2)

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.

Lateral Economics also highlighted the effects of ethics approval on the evaluation of the Family by Family program.

... numerous preliminary ethics processes took around nine months, though this was simply to ask families questions about their progress — as they'd been asked regularly within the program. The evaluation ignored the program's effect on children. Why? Because getting that aspect

through the ethics procedures would have been too expensive, uncertain and time-consuming. How ethical can you get? (sub. DR177, pp. 13–14)

In this example, Lateral Economics argued that the ethics process got in the way of being able to capture and demonstrate the success and innovations involved with this program. Other feedback was about the considerable cost of undertaking research across states and territories, with different ethics requirements and review processes.

Getting the balance right is a challenge. In the draft Background Paper, the Commission referred to relevant guidance on when a formal ethics review process may not be required for evaluation, including when 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).

However, many participants were concerned that agencies should not rely on hard rules to make ethics assessments. Yulang Indigenous Evaluation noted that ‘there are many situations where an evaluation relying exclusively on existing data would require formal ethical review and approval. In fact, the NHMRC’s Ethical Considerations in Quality Assurance and Evaluation Activities specifically warns that secondary use of data is a trigger for considering that formal ethical review may be necessary’ (sub. DR150, p. 15).

The CSIRO also commented that:

Care should be taken when excluding ... projects from ethics review as they may introduce significant risks to evaluations of programs and policies. Evaluations using de-identified data or existing datasets may still have an impact on Aboriginal and Torres Strait Islander peoples and communities, and ethical review is essential in planning for and navigating the conduct of the evaluation and ensuring it is done so according to the principles of the strategy. (sub. DR131, p. 10)

And NSW Aboriginal Housing Office said:

... the use of existing data, even when de-identified, is open to the risk of perpetuating deficit discourses about Aboriginal people as it is analysed and interpreted in the absence of other contributing external, contextual factors. For this reason, the ... statement ... [in the draft Strategy regarding when a formal ethics review process may not be required] ... potentially overlooks the need to assess this on a case-by-case basis and potentially signals a green light to evaluators in situations where there is a need for ethics review and approval. (sub. DR130, p. 2)

The Independent Members of the National Indigenous Australians Agency Indigenous Evaluation Committee also cautioned against the idea that using existing or de-identified data would require lesser ethical review:

As demonstrated in the growing literature in Australia and elsewhere on the problematics of Indigenous data, especially in relation to Open Data, there remains a strong potential for existing data and findings to be used in ways that adversely impact Aboriginal and Torres Strait Islander Peoples. Moreover, ethical review processes require a substantial investment in time and effort and the risk is that departments and agencies may be tempted to design their evaluations to avoid such reviews if possible. This could have a detrimental effect on evaluation quality. (sub. DR115, p. 4)

Government agencies therefore need to carefully consider how ethics assessments are made to ensure that all evaluations are undertaken ethically, even those evaluations where a formal ethics review by an ethics committee is not pursued. It is important that government agencies clearly document how they will assess ethical risks, and make decisions on whether to request formal review by an ethics committee — making sure they are consistent with the existing research guidelines referred to above.

Decisions on the necessity of a formal ethics review need to be transparent, and reflect Aboriginal and Torres Strait Islander perspectives. This highlights the need for agencies to establish appropriate governance arrangements, where Aboriginal and Torres Strait Islander people share in decision making around how ethical risks are assessed. The Victorian Aboriginal Child Care Agency, for example, suggested that the existing HREC processes may not be adequate to ensure the ethics review process centres Aboriginal and Torres Strait Islander people, noting that ‘... approval should be by an HREC with an Aboriginal and Torres Strait Islander researcher member ... While not all HRECs currently include such members, ... the onus is on HRECs to address this, or at a minimum, undertake review and consultation with an Aboriginal or Torres Strait Islander researcher as appropriate’ (sub. DR126, p. 3).

Consistent with other quality standards referred to in this chapter, taking an ethical approach also requires ethical processes being integrated into early planning and policy design so that sufficient time and resources are dedicated to evaluations. For example, if ethics is only considered after the evaluation has been designed and commissioned, then it may already be too late to meet the requirements of an ethics review.

6 What to evaluate

Key points

- It is not practical or feasible to evaluate all policies and programs. Setting priorities ensures the limited resources Australian Government agencies and Aboriginal and Torres Strait Islander people, communities and organisations have are directed to evaluations that can best inform the development of policies and programs to improve the lives of Aboriginal and Torres Strait Islander people.
- Currently there is no systematic whole-of-government approach to prioritising evaluations. 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.
- The National Agreement on Closing the Gap provides the basis to inform an interim set of priority evaluations aimed at improving the lives of Aboriginal and Torres Strait Islander people. These will guide agencies' efforts on what to evaluate under the Indigenous Evaluation Strategy — that is, beyond agency-specific considerations. These priority areas should be formalised by the proposed Office of Indigenous Policy Evaluation in partnership with the proposed Indigenous Evaluation Council.
- A criteria-based priority setting process should be adopted by each Australian Government agency to determine how its new and existing policies and programs that impact the lives of Aboriginal and Torres Strait Islander people should be prioritised for evaluation. 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 determined 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. It is also not practical nor feasible to evaluate 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 priorities 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 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 (discussed below).

A priority-setting framework can also improve accountability in decision making by ensuring there is a systematic and transparent basis for determining and resourcing evaluations. This means Aboriginal and Torres Strait Islander people, 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 not get evaluated. This is particularly important for mainstream policies and programs that can have significant impacts on Aboriginal and Torres Strait Islander people. A priority-setting 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, and ways to improve how they work together to achieve whole-of-government priority outcomes.

Balancing the costs with the benefits of 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 evaluation, it is useful to think about the ‘opportunity cost’ of an evaluation (or how else the resources used for evaluation could be used). This is because 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 participants to inform the evaluation (chapter 5). 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 amount of engagement undertaken (including users, providers and communities) — particularly where meaningful engagement involves Aboriginal and Torres Strait Islander people leading or partnering in the evaluation process.

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).³⁴ This is in addition to the costs these evaluations place on government and also on the communities engaged to participate in them.

³⁴ 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’, and ‘conference’. Also excluded the category of ‘individual health screening and assessment services’.

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. 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)

The costs of evaluation on Aboriginal and Torres Strait Islander people and organisations, in light of the benefits, was raised by a number of participants. Indigenous Allied Health Australia, for example, said:

It is important that direct, outcome focussed benefits achieved through evaluation processes exceeds the additional burden which may be placed on service providers, particularly in the community-controlled sector. (sub. 31, p. 4)

NSW Aboriginal Land Council also noted:

... without stronger commitments on how evaluations will be used by Governments, there is a risk the Indigenous Evaluation Strategy will add to the burden of reporting many Aboriginal Community Controlled organisations currently face, without delivering any benefits. (sub. DR122, p. 3)

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.

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). Academics from the Centre for Aboriginal Economic Policy Research at the Australian National University also suggested that there is a risk the Strategy 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 strategic approach to priority-setting should evolve over time, as government agencies increasingly put Aboriginal and Torres Strait Islander people and communities at the centre of policy and program evaluation and as emerging successes spread with new knowledge of what's working in the field (Lateral Economics, sub. DR177, p. 8).

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 exist 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, these processes 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 through its renewed commitments under the National Agreement on Closing the Gap (the National Agreement), which came into effect in July 2020.

As discussed in chapter 1, the National Agreement reflects the collective commitment of all Australian governments (national, state, territory and local government), and was developed through a formal partnership with Aboriginal and Torres Strait Islander people (represented by the Coalition of Aboriginal and Torres Strait Islander Peak Organisations). This has been a significant shift from previous efforts where a review by the Australian National Audit Office found there was limited consultation with Aboriginal and Torres Strait Islander people in the development of the previous Closing the Gap framework (ANAO 2019a, p. 30).

In developing its implementation actions under the National Agreement, the Australian Government has committed to be guided by ‘ … partnering with and seeking the views and expertise of Indigenous Australians, including Elders, Traditional Owners and Native Title holders, leaders, communities and organisations … ’ (NIAA 2020a).

The focus of the National Agreement is on a set of priority reforms and socio-economic outcomes to be pursued by governments and Aboriginal and Torres Strait Islander people, it also includes review arrangements to track progress against these commitments. The Commission proposes that evaluation efforts under the Strategy should largely support activities being pursued through the National Agreement by prioritising evaluations of these efforts as a starting base (section 6.3).

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)

Formal evaluation priority-setting processes within agencies

Based on responses to the information request the Commission sent to Australian government agencies, only around 35 per cent reported having a formal criteria for determining what policies and programs are evaluated (appendix B, figure B.8). In general, agencies that have a formal framework in place for setting evaluation priorities take a proportional, principles-based approach to deciding what (and to what extent) a policy or program is evaluated (box 6.1). The principles-based approach is generally based on the size (as measured by impact or expenditure), strategic significance or risk profile of the policy or program (appendix B, figure B.9).

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)

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 2017, p. 15).

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 assist 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 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 (2019c, pp. 7, 33, 38); Department of the Environment (2015); DFAT (2017a, 2017b, p. 4); DIIS (2017); DPMC (2018).

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 rarely 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.

6.3 Priority-setting under the 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 Aboriginal and Torres Strait Islander communities 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, to align with the overarching principle of the Strategy, 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 the National Agreement on Closing the Gap

The framework of priority reforms and socio-economic outcomes identified in the National Agreement should shape government-wide evaluation priorities. The National Agreement:

... is a pledge from all governments to fundamentally change the way they work with Aboriginal and Torres Strait Islander communities and organisations ... Priority Reforms commit governments to new partnerships with Aboriginal and Torres Strait Islander communities across the country; strengthen community-controlled organisations to deliver closing the gap services; address structural racism within government agencies and organisations; and improve sharing and information with Aboriginal and Torres Strait Islander organisations to support shared decision making. ... The National Agreement also establishes 16 national socio-economic targets in areas including education, employment, health and wellbeing, justice, safety, housing, land and waters, and Aboriginal and Torres Strait Islander languages. The targets bring focus to new areas important to the lives of Aboriginal and Torres Strait Islander people and will help to monitor progress in improving their life outcomes. (Coalition of Peaks 2020a)

Aligning evaluation priorities with the National Agreement was endorsed by a number of participants.³⁵ For example, ANTaR said:

... the [Indigenous Evaluation Strategy] will necessarily focus on the programs and services that will be prioritised under the next Closing the Gap Framework to be agreed between the Australian governments and Aboriginal and Torres Strait Islander peoples through the Coalition of Peaks. The Federal Government's commitment to evaluation through the establishment of the [Indigenous Evaluation Strategy] should become a key plank of the Closing the Gap agreement and push the evolution of the Framework over the coming decade. (sub DR124, p. 6)

While the Commission has aligned the Strategy to the policy framework defined by the National Agreement, it is an Australian Government Strategy, aimed at improving evaluation culture across Australian Government agencies. It is possible for this relationship with the National Agreement to be strengthen over time as governments settle on how they intend implementing their responsibilities under the National Agreement. State and territory governments, for example, may see value in adopting the principles and framework of the Strategy to monitor their progress.

Defining evaluation priorities from policy priorities

The policy framework under the National Agreement is not focused on the evaluation of specific actions that will be required to achieve its objectives. A process to transform these priority reforms and socio-economic outcomes into a set of priority government-wide evaluations is required (figure 6.1). This process should also consider other evaluation priorities aimed at improving the lives of Aboriginal and Torres Strait Islander people that are not captured under the National Agreement, 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:

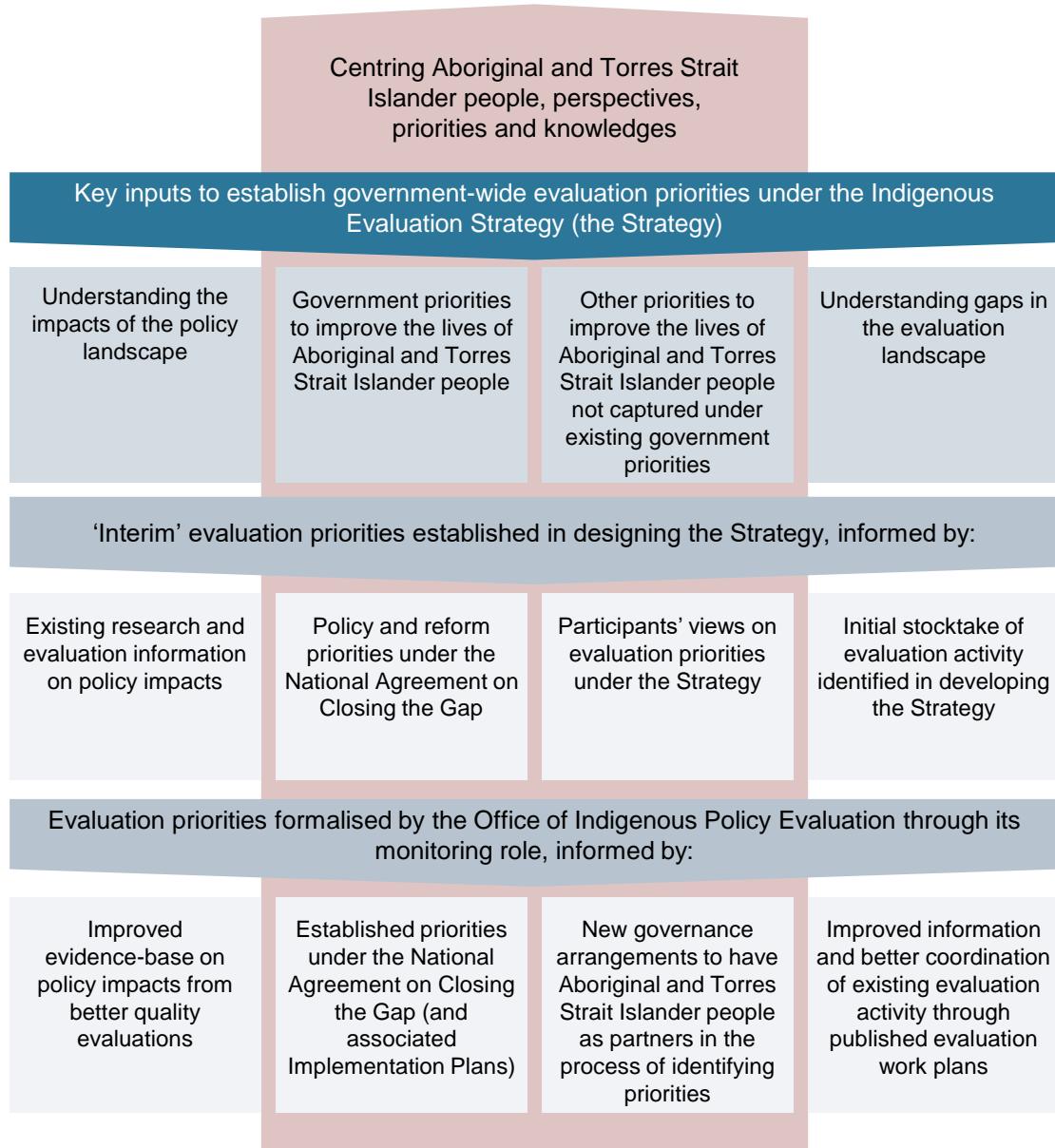
- scope of government activity in these priority areas, such as expenditure, the number, size and scope of programs, and the prominence of these programs in the government's agenda
- potential impact of activities (policies or programs) in these priority areas
- current evaluation activity (and importantly the gaps) in these priority areas.

The task of establishing government-wide evaluation priorities would initially sit with the proposed independent Office of Indigenous Policy Evaluation (OIPE), and then the proposed Centre of Evaluation Excellence if it is established. In line with the overarching principle of

³⁵ Aligning evaluation priorities with the National Agreement was endorsed by AIHW (sub. DR123, p. 3); ANTaR (sub. DR124, p. 6); Community First Development (sub. DR134, p. 9); Empowered Communities (sub. 41, p. 11); NACCHO (sub. DR170, p. 5); NATSIHSC (sub. DR148, p. 3); NIAA (sub. DR161, p. 3); NSW Aboriginal Housing Office (sub. DR130, p. 1); QAIHC (sub. DR173, p. 2); the Royal Australasian College of Physicians (sub. 77, p. 6; sub. DR138, p. 6); the Tasmanian Government (sub. 100, p. 5); and the Western Australian Government (sub. 74, p. 10).

the Strategy (and priority reform one under the National Agreement), the proposed Indigenous Evaluation Council (with all Aboriginal and Torres Strait Islander members) would partner with the OIPE to establish these evaluation priorities under the Strategy (chapter 7).

Figure 6.1 Process for formalising government-wide evaluation priorities



An interim set of government-wide evaluation priorities

The Commission was asked to identify evaluation priorities. However, setting evaluation priorities without having developed partnership arrangements with Aboriginal and Torres Strait Islander people would be inconsistent with the overarching principle of the Strategy of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges.

As such, the Commission has identified an ‘interim’ set of evaluation priorities based on the outcomes of the National Agreement, and:

- 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.

On the latter point, the Commission found that, 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 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 the 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 Commission undertook more targeted engagement with Aboriginal and Torres Strait Islander people between the draft and final Strategy on what the evaluation priorities should be. This included a dedicated workshop with Aboriginal and Torres Strait Islander leaders from a number of Aboriginal and Torres Strait Islander peak organisations, government organisations, and academic institutions, as well as non-Indigenous leaders from key government agencies (appendix A).

Specific policies and programs that would benefit from evaluation across the eight policy domains identified in the National Agreement (chapter 3, section 3.3) were identified. But most of the workshop feedback was on how the Strategy could promote evaluation efforts across the priority reforms under the National Agreement (participants saw these as key to achieving the aims of the National Agreement, including its socio-economic outcomes and targets). There was also some support for including cross system priorities, which had earlier been identified as part of the approach to developing and implementing the next phase of Closing the Gap, namely to identifying actions across multiple targets and policies to:

... address intergenerational change, racism, discrimination and social inclusion (including in relation to disability, gender and LGBTIQ+), healing and trauma, and the promotion of culture and language for Aboriginal and Torres Strait Islander peoples. (COAG 2018, p. 2)

The intent of the cross-system priorities is reflected in the new priority reforms, particularly priority reform three, which focuses on ‘systemic and structural transformation of mainstream government organisations to improve accountability and respond to the needs of Aboriginal and Torres Strait Islander people’ (JCOCTG 2020, p. 11). Under this priority reform, governments committed to: identify and eliminate racism; embed and practice meaningful cultural safety; deliver services in partnership with Aboriginal and Torres Strait Islander organisations, communities and people; increase accountability through transparent funding allocations; support Aboriginal and Torres Strait Islander cultures; and improve engagement with Aboriginal and Torres Strait Islander people.

The other strong theme from the feedback was that the governance arrangements for setting evaluation priorities should reflect the principles of Aboriginal and Torres Strait Islander leadership, as well as of partnership and shared decision making with Aboriginal and Torres Strait Islander people.

The governance arrangements proposed for the Strategy (and the priority-setting process) have Aboriginal and Torres Strait Islander leadership and shared decision-making via the Indigenous Evaluation Council (chapter 7). As discussed later, the way agencies decide on agency-level priorities will also need to reflect the centring of Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. As noted by Yulang Indigenous Evaluation, this will need to be resourced appropriately to ensure ‘diversity of Aboriginal and Torres Strait Islander peoples, and involvement of community members’ (sub. DR150, p. 17).

The draft Strategy’s proposed government-wide evaluation priorities focused on the seven ‘priority policy areas’ included in early engagement documents by the Coalition of Aboriginal and Torres Strait Islander Peak Organisations (2018, 2020). The list captured 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. It also reflected some of the key expenditure areas identified in the *Indigenous Expenditure Report* (with the omission of social security and welfare support programs) (chapter 3).

The National Agreement sets out 16 socio-economic outcomes across eight policy domains. These largely reflect additional and continuing efforts across the previously defined ‘priority policy areas’, with the main change being the addition of outcomes in a new ‘culture and languages’ policy domain. It is proposed that effort across these policy domains continue to be a priority for evaluation activity under the Strategy (figure 6.2).

Reflecting the feedback and the National Agreement, it is proposed that the four priority reforms, rather than sitting alongside the priority policy domains, should sit above them. This would help shape evaluation priorities within (and across) these policy domains and inform the way evaluations should be conducted. Therefore evaluations under the Strategy should prioritise:

- **New or existing partnership or shared decision-making arrangements** involving the Australian Government with Aboriginal and Torres Strait Islander people. These evaluations would focus on their effectiveness at meeting objectives, including the partnership elements listed in clauses 32 and 33 of the National Agreement. Under this stream of evaluations, Australian Government agencies (either independently or jointly with other parties) would evaluate joined up approaches in the areas of justice (adult and youth incarceration), social and emotional wellbeing (mental health), housing, early childhood care and development, and Aboriginal and Torres Strait Islander languages (as identified in clause 38 of the National Agreement).
- **New or existing efforts by Australian Government agencies to build a strong Aboriginal and Torres Strait Islander community-controlled sector.** Under this stream of evaluations, Australian Government agencies (either independently or jointly with other parties) would evaluate joint strengthening efforts in the early childhood care and development, housing, health, and disability sectors (as identified in clause 50 of the National Agreement).
- **New or existing efforts to transform mainstream government organisations (including their policies and programs), to be more accountable and responsive to the needs of Aboriginal and Torres Strait Islander people.** Focusing on the high impact activities of Australian Government agencies across the priority policy domains, evaluate Australian Government agency efforts to deliver these policies and programs in a way that: eliminates racism; embeds and practices cultural safety; delivers services in partnership with Aboriginal and Torres Strait Islander organisations, communities and people; increases transparency in funding allocation; supports Aboriginal and Torres Strait Islander cultures; and improves engagement with Aboriginal and Torres Strait Islander people. There is a substantial number of specific policies and programs identified by participants in this study that could be a focus for policy reform and/or evaluation under this stream (chapter 3, section 3.3).
- **New or existing efforts by Australian Government agencies to share access to location specific data and information with Aboriginal and Torres Strait Islander communities and organisations.** Under this stream, Australian Government agencies would evaluate the success of data projects established in up to six locations across Australia to enable Aboriginal and Torres Strait Islander communities and organisations to access and use location-specific data on the Closing the Gap outcome areas (as per clauses 74 and 75 in the National Agreement).

The priorities are focused on the National Agreement's priority reforms and policy domains, which align with the Strategy's objective of improving the lives of Aboriginal and Torres

Strait Islander people, rather than focus on the more narrowly defined indicators (or targets) of ‘closing the gap’. As the National Justice Project noted:

... 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 Aboriginal and Torres Strait Islander Peak Organisations, there is broader support around the policy priority outcomes (relative to the targets) amongst Aboriginal and Torres Strait Islander people (Coalition of Aboriginal and Torres Strait Islander Peak Organisations 2020, p. 7).

Figure 6.2 Proposed interim government-wide evaluation priorities
Based on the National Agreement on Closing the Gap

The Strategy should prioritise evaluation of government efforts ...			
... delivered as part of a formal partnership and/or shared decision making arrangement (Priority 1) that build up the Aboriginal and Torres Strait Islander community controlled sector (Priority 2) (particularly mainstream ones) that eliminate racism, embed and practice meaningful cultural safety, deliver services in partnership with Aboriginal and Torres Strait Islander people (Priority 3) involved in data sharing arrangements (Priority 4) ...
... and focused across (and within) the following policy domains :			
<ul style="list-style-type: none">• Education• Housing• Health and wellbeing	<ul style="list-style-type: none">• Land and water• Justice• Culture and languages	<ul style="list-style-type: none">• Safety (Families, children and youth)• Employment (Economic development)	

Source: adapted from Joint Council on Closing the Gap (2020).

Prioritisation process within agencies

Identifying scope of policies and programs for evaluation: Flows and Stock

The prioritisation process is about knowing what policies and programs are in place across Australian Government agencies that align with the set of government-wide evaluation

priorities (figure 6.2), and that should be considered for evaluation (under the process illustrated in figure 6.3).

It is the responsibility of all government agencies, as stewards of the policies and programs they administer, to identify those that contribute to government-wide evaluation priorities — along with any 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 (Strategy action 1). 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 mainstream policies and programs that have the largest impacts on the lives of Aboriginal and Torres Strait Islander people, and where there would be value in focusing evaluation efforts under the Strategy.

In short, the prioritisation process within agencies will ensure the process of selecting what to evaluate is more systematic; reflects the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people; and is transparent.

Figure 6.3 Priority setting under the Strategy



For new policies and programs, agencies would be required to identify if the activity proposed for Cabinet or Government (in the case of decisions by Ministers and/or statutory authorities) endorsement will significantly impact the lives of Aboriginal and Torres Strait Islander people (and/or contribute to any of the government-wide evaluation priorities). If so, it should be considered in scope for evaluative activity under the Strategy. This

requirement on agencies to embed early impact assessment (not to be confused with impact evaluations) and evaluation planning is proposed in chapter 5 (Strategy action 3).

For the stock of policies and programs, agencies should undertake a stocktake of existing policies and programs they administer to determine whether they are in scope of the Strategy. The Commission has not specified a process for undertaking a stocktake — agencies should decide on an approach based on the nature and scope of their policies and programs as they relate to the Strategy. 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 areas, for example, through sunsetting 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.

The Minister for Indigenous Australians is leading the development of the Australian Government's implementation plan under the National Agreement. It will:

... set out how policies and programs are aligned to the National Agreement and what actions will be taken to achieve the Priority Reforms. They will also include information on funding and timeframes for actions. (NIAA 2020a)

This is an opportunity for agencies to provide a bottom up assessment of the actions it is taking, or could take, to support the priorities of the National Agreement, and to plan out an evaluation strategy for these activities.

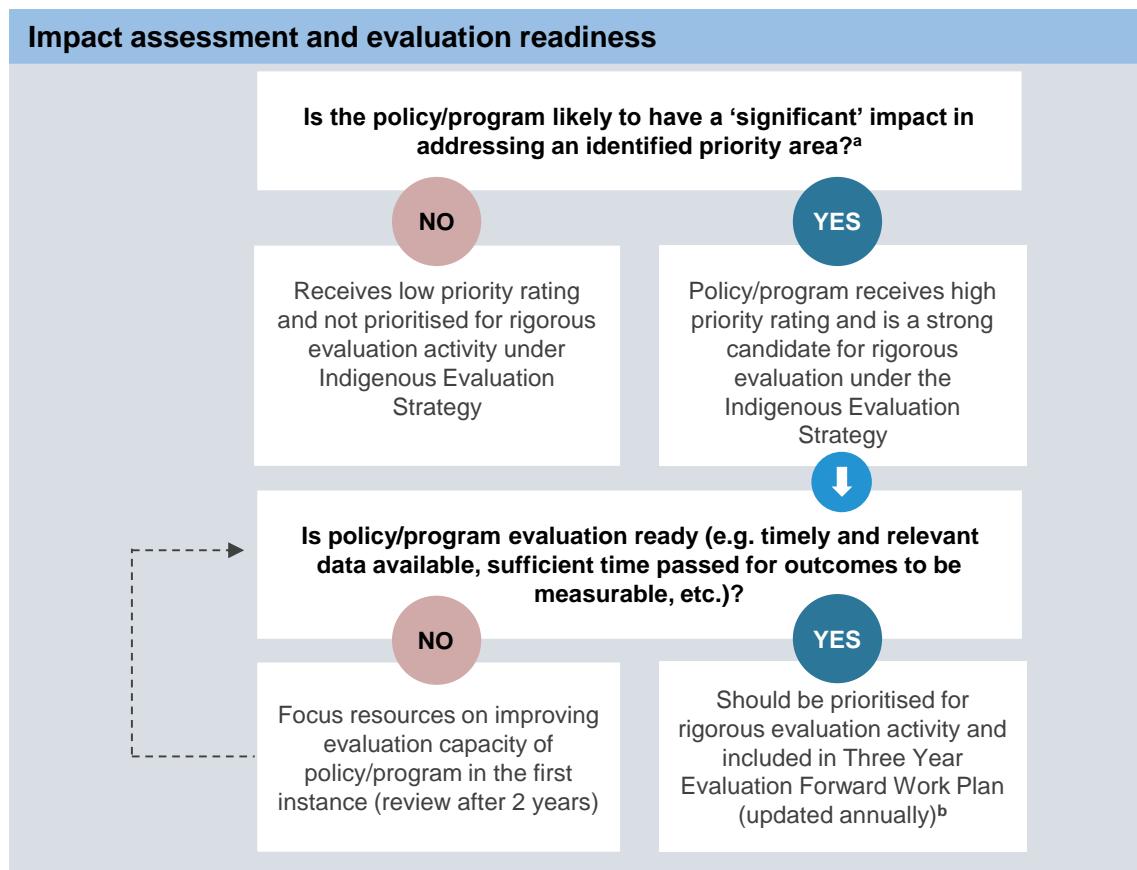
The stocktakes should be transparent, including publishing in Three Year Evaluation Forward Work Plans (discussed below), details on:

- 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.

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. Under the Strategy, it is proposed that this be done with the aid of impact and evaluation readiness assessments (figure 6.4).

Figure 6.4 Prioritising evaluations under the 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.

There is a fairly standard set of impact assessment criteria used for deciding what to evaluate. These are summarised in table 6.1. While participants had different views on what criteria should be used to determine key priorities for evaluation (box 6.2), most suggested criteria commonly used in priority-setting.

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, this

assessment should reflect the risks as they relate to Aboriginal and Torres Strait Islander people (such as the risk of a program with limited evidence of working in Aboriginal and Torres Strait Islander communities). This highlights the importance of making sure that those making these assessments have the skills and knowledges required to reflect the needs of Aboriginal and Torres Strait Islander people. This is why it is important for agencies to establish 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 impact assessments.

Table 6.1 Impact assessment criteria for priority-setting

<i>Criterion</i>	<i>Summary</i>
Impact of program	This principle is based on how the program will likely 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.

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.

Agencies should develop a fit-for-purpose multi-criteria impact 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 (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 impact 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 (Guide, section 2, box 3).

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, 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).

Government agencies need to weigh up multiple 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 high, it could be prioritised for evaluation. Alternatively, an established program with large expenditure but that has been recently evaluated, may be less of a priority for formal evaluation activity in the short term.

In establishing and maintaining the impact assessment criteria, agencies should engage Aboriginal and Torres Strait Islander people to ensure the metrics used in these assessments reflect the perspectives, priorities, and knowledges of Aboriginal and Torres Strait Islander people. Establishing strong partnerships to support agency evaluation efforts will take time and resources, and some agencies will need to rely on other engagement methods to develop an initial set of impact assessment criteria — which could involve a process of peer-review or transparent consultation mechanisms with agency stakeholders where feedback is sought and responded to.

To ensure agencies use these criteria effectively, there are a few requirements that they should be expected to follow:

1. impact assessments should centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges. This requires agencies to have appropriate governance to ensure Aboriginal and Torres Strait Islander people are engaged in making these decisions (box 6.3)
2. the impact assessment methodology should be publicly available (and included as part of each 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 impact assessment methodology should be reviewed and updated over time to ensure it remains relevant (section 6.4).

Box 6.3 Centring Aboriginal and Torres Strait Islander people, in 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 inform, and make decisions about, 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, ... [we consider] ... 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, other policies and programs should still be developing their own evidence base. Agencies should develop thresholds to further categorise lower levels of evaluative activity for those policies and programs 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 will facilitate learning by doing (such as user surveys or reviews). This will ensure that evaluation culture is embedded in policy design and implementation across government, and also that these programs will be ready for more rigorous evaluation activity in the future, if required.

There may also be other evaluations set outside the Strategy, including sunset clauses and reviews of major programs that are set when policies and 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, evaluation timing 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 the most appropriate timing. 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).

The former Department of Industry, Innovation and Science *Evaluation Ready* process is a tool for considering what is required for evaluation readiness (box 6.4). 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 2019a, pp. 12–17).

In the early days of the Strategy’s implementation, it may not be feasible for all significant policies and programs to be rigorously evaluated. This is not to suggest that no activity should be initiated for these policies and programs, rather that 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**³⁶ 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 (2017).

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 proposes (Strategy action 2) that each department centrally coordinate, on an annual basis, a rolling Three Year Evaluation Forward Work Plan, which would detail:

- 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
- how the department identified high priority policies and programs
- 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)

³⁶ 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.

-
- 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?

It is proposed that Plans would be publicly released on the department's website, and provided to the OIPE so it could 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 decide whether larger agencies within their portfolio 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.

Departments could use the development of the Three Year Evaluation Forward Work Plans as an opportunity to feed up potential actions to support the Australian Government's implementation planning and ongoing reporting as part of the National Agreement.

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 endorsed, it is proposed that departments and major agencies would 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 would have developed an initial rolling Three Year Evaluation Forward Work Plan. The initial plan would unlikely 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 principles-based, and largely devolved framework, the priority-setting process should align with existing processes within agencies. Agencies would not be required to adopt all assessment criteria outlined in the Strategy if it has made an assessment that its existing criteria adequately cover 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 proposed 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 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 by what works, in practice, evaluation findings are often not used to drive policy or program design, or to inform agencies about where to invest taxpayer funds.
- Evaluations are more likely to be used if they address the questions that decision makers and those affected by the policy or program (in this context, Aboriginal and Torres Strait Islander people, organisations and communities) want answered. This points to the importance of embedding evaluation into policy making processes.
- Policy makers need to know about evaluation evidence to use it. Evaluation findings should be published wherever possible. Transparency of evaluation results can also improve the quality of evaluations, facilitate a 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 to use. A central clearinghouse for evidence on the effectiveness of policies and programs affecting Aboriginal and Torres Strait Islander people would strengthen accountability for undertaking good evaluations and centring Aboriginal and Torres Strait Islander people, and improve the diffusion of knowledge.
- External oversight would strengthen accountability and help to lift the bar on the quality and usefulness of evaluations. An Office of Indigenous Policy Evaluation (OIPE), housed within an existing Australian Government agency, should monitor and report on agencies' progress in implementing 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 foreground Aboriginal and Torres Strait Islander people. An Indigenous Evaluation Council should work in partnership with the OIPE to identify Australian Government-wide evaluation priorities and report on agencies' implementation of the Strategy.
- 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, an 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 and the Indigenous Evaluation Council could move to the Centre as a standalone branch.

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 (the 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. (2020c, p. 63)

The Commission consulted with many Australian Government agencies while developing the 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* 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 across the APS — such a culture will be critical for the Strategy to be effective.

- 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.

Box 7.1 A culture of evaluation and learning is one that values and prioritises 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 policy making is something governments should practice, including several iterations from the Productivity Commission: 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’ (Stewart, 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.

Mayne suggested that an organisation with a strong culture of evaluation:

- engages in self-reflection and self-examination (that is, 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).

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).

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 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. The Australian Government's modest response to the recommendations of the Independent Review of the APS aimed at building evaluative culture and capability (chapter 2) means that the Strategy will need to create incentives for agencies to develop a culture of evaluation and learning.

Organisational culture change will not happen overnight. As Te Arawhiti (the Office for Māori Crown Relations) acknowledged in the Organisational Capability component of its *Māori Crown Relations Capability Framework* for the New Zealand Public Service:

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)

Are there barriers to developing a culture of evaluation and learning?

While there is strong agreement across Australian Government agencies that policy should be informed by what works, in practice there are several factors (often related) that can work against the systematic generation and use of evidence for policy and program design.

Evaluation can — and should — support both accountability and learning (chapter 1). But these two goals can create competing incentives for agencies: while a culture of learning from evaluation requires some tolerance for experimentation and policy mistakes, the spending of public funds (and governments' responsibility to provide effective public services) typically demands consequences for poor policy design and implementation.

Many participants expressed a concern that agencies' evaluation activity is often more focused on accountability and compliance (for example, by focusing on how funds are spent and how many people access services) than on innovation and learning (even when, in many policy areas affecting Aboriginal and Torres Strait Islander people, the solutions are not known to governments). 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)

A recent ANZSOG paper, prepared for the Independent Review of the APS, explained that the nature of policy making today 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)

The demands of the everyday work of government, and the realities of politics (including extensive media coverage), mean that agencies have incentives to be risk-averse rather than maximising opportunities to learn from innovation and experimentation — there can be harsh consequences from getting policies wrong (for example, DoFD 2012, p. 6; Luetjens and ‘t Hart 2019, p. 7; Shergold 2015, p. 81). 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)

Time limitations, political pressures and competition for scarce resources may also drive policy makers to focus on achieving results quickly, at the expense of trying to understand why particular problems exist and why particular solutions might or might not work (Australian Council of TESOL Associations, sub. DR179, pp. 13–15; Luetjens and ‘t Hart 2019, p. 10).

In such an environment it is easy to see why evaluation (and learning from success and failure overall) may take a back seat. There is no panacea: the challenge is to evolve systems, attitudes and capabilities in the public sector so that scrutiny of policies and programs is seen as integral to government activity, rather than as something that ‘happens to’ agencies (ANAO 2014b, p. 12).

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)

Some participants suggested that even where evidence on the effectiveness of policies and programs that affect Aboriginal and Torres Strait Islander people exists, it is often ignored or used selectively (Cape York Institute, sub. 69, p. 7; Shelley Bielefeld, sub. 65, p. 2).

Results of a recent OECD survey also suggested that, internationally, 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 2020c, p. 22).

The question is, given the realities of policy making, what can be done to boost the ‘demand side’ of the equation — the extent to which government agencies value and use evaluation evidence 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 utilisation-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). Other potential research users — such as service providers or advocacy bodies — are also 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 (discussed in chapter 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:

... intended 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)

Box 7.2 Utilisation-focused evaluations

Michael Quinn Patton's concept of utilisation-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 utilisation of both the findings and of the evaluation process itself to inform decisions and improve performance.

Utilisation-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 utilisation-focused evaluation.
2. Assess and enhance evaluator readiness and competence to undertake a utilisation-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. Organise 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).

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.

All of these factors point to the need for evaluation to be integrated into the policy-making process (chapter 5). This, in turn, requires much of the work of planning, commissioning and

conducting evaluations to be embedded within the agencies that design and deliver programs and policies.

Participants' views on this aspect of evaluation practice were mixed — some argued that evaluations should be commissioned and conducted at arms' length from policy and program agencies, for independence and objectivity,³⁷ while others said that evaluation needs to become better integrated into the policy cycle (NSW Ombudsman, sub. DR152, pp. 3–4), and that individual agencies should develop ongoing evaluation partnerships with Aboriginal and Torres Strait Islander organisations and communities (Empowered Communities, sub. DR171, pp. 5–6) (implying that evaluation needs to take place at the agency level).

As discussed in chapter 5, integrating evaluation into the policy cycle is key to quality evaluations. As such, agencies should be responsible for managing evaluations of the policies and programs that they design and implement. An appropriate level of independence can be achieved by requiring evaluations to be conducted by evaluators that are independent of the area of the agency with policy or program responsibility. And external oversight (section 7.3; chapter 10) can provide an independent quality assurance function.

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. (2020c, 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. For example, 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

³⁷ For example: Tony Dreise et al (sub. 33, pp. 5–6); Victorian Aboriginal Community Controlled Health Organisation (sub. 44, p. 20); Australian Council of TESOL Associations (sub. 87, p. 41).

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). And even where agencies have such a policy, the publication of evaluations can be delayed because of Ministers needing to note the completed evaluation report, or because there are sensitivities around the findings and/or material contained in reports (for example, ANAO 2019b, p. 46).

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 accountability, and to 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.

As well as increasing the use of evaluation results in policy formulation and modification, requiring evaluation results to be made public in a timely fashion can:

- allow for external scrutiny of evaluations’ approaches, methodologies, processes and analyses, and provide an incentive to improve the quality of evaluations
- 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.

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)

Jobs Australia:

There should be a presumption that all evaluation reports are made publicly available in a timely fashion. Aside from providing assurance and confidence in the integrity of government, the publication of evaluations ensures that learnings can be readily picked up by relevant stakeholders and more readily adapted. ... In [some] cases, a modified version that protects program participants may need to be developed which both captures the learnings and retains participant anonymity. (sub. 57, pp. 8–9)

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. 15)

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)

Kimberley Aboriginal Medical Services:

... simply generating great quality evaluations is not enough. The space for sharing of these learnings, the receptiveness of government and services to heed the lessons learnt is critical. (sub. DR164, p. 2)

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 2020c, 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). The results of our information request suggest that, in the Australian Government, only about 8 per cent of responding agencies always made evaluations public (chapter 3, appendix B).

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, organisations and communities is particularly important in light of Australia's colonial history and its continuing impacts (chapter 1), including the harmful impacts of research and evaluation in Aboriginal and Torres Strait Islander communities (Katz et al. 2016, p. 36, referenced in Yulang Indigenous Evaluation, sub. DR150, p. 9).

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 Strategy (Strategy, action 7). The risk that this weakens agencies' incentives to undertake evaluations, or motivates agencies to influence or 'sanitise' the content of evaluation reports, would be significantly mitigated by a transparent process of setting and tracking evaluation priorities (chapter 6) and an independent oversight function with a mandate to assess the quality of evaluation reports and report to the public (section 7.3, chapter 10).

Where there are compelling concerns that publishing a full evaluation report would compromise confidentiality or privacy, or where there is culturally sensitive information, a summary report should be published instead (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 (Strategy, action 10; discussed in more detail in the next section).

Agencies' responses to evaluation findings

The way evaluations lead to better outcomes is via agencies implementing changes in response to lessons learned from evaluations. A number of participants suggested that agencies should be required to formally respond to evaluation findings (this is the case in Canada — appendix D).³⁸

³⁸ Requiring agencies to respond to evaluation findings was suggested by: Victorian Aboriginal Community Controlled Health Organisation (sub. 44, p. 22); Jobs Australia (sub. 57, p. 6); NSW Aboriginal Land Council (sub. DR122, p. 3); The Fred Hollows Foundation (sub. DR156, p. 5); National Voice for Our Children (SNAICC) (sub. DR163, p. 14).

There are a number of benefits to agencies formally responding to evaluation findings, including:

- demonstrating that the agency values (and uses) evaluation evidence for improving policies and programs
- allowing other evaluation users (agencies, Aboriginal and Torres Strait Islander organisations, service providers and others) time to incorporate any necessary changes to the policy or program into their planning
- allowing the public, including Aboriginal and Torres Strait Islander people, organisations and communities, to track agencies' progress implementing those changes and hold them accountable.

Given that a minority of all evaluations for Australian Government agencies are published, it is also rare for agencies to publish responses to evaluation reports. The Office for Development Effectiveness (ODE), within the Department of Foreign Affairs and Trade (DFAT), was one exception.³⁹ Under the 2017 *Aid Evaluation Policy*, all management responses to the ODE's strategic evaluations were to be finalised within 28 days of completion of the evaluation report. The completed report, along with the corresponding management response, was then published on the Department's website (DFAT 2017a, p. 3).

The Strategy (action 9) proposes that agencies' responses to evaluation conclusions should include: an explanation about what they have learned, what they have changed in response to the findings/recommendations, and any further action they intend to take on the relevant policies or programs. These responses could be included as part of the evaluation report, or published separately.

Including 'lessons learned' in evaluation reports, and processes for ensuring that evaluation results feed into future policymaking, are both discussed in more detail below. Agencies' tracking of their own planned actions in response to evaluation conclusions could also be incorporated into their reporting on evaluation progress against Three-Year Evaluation Forward Work Plans (chapter 6).

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

³⁹ Structural changes in September 2020 to the way DFAT approaches evaluation of its aid and development work mean that the ODE no longer exists as an independent unit (and with an Independent Evaluation Committee). It has been folded into the Office of the Chief Economist (Cornish 2020). It is not yet clear whether this will result in changes to the *Aid Evaluation Policy* in the near future.

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 funded from 2009 to 2014: chapter 2) — 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 evaluation evidence

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)

Institute for Human Security and Social Change:

... evaluation which is cross-cultural and where different worldviews and values are likely to encounter each other, needs to carefully consider the analysis processes, which build dialogue and connection between different worldviews. While this may not be a simple process, it is if handled well it is likely to lead to more active attention to solving problems and working together for improvement. If the Strategy is serious about wanting evaluation evidence to influence programming in practice, then it should promote good practice for analysis, in particular analysis which supports active critical dialogue between stakeholders. (sub. DR119, p. 3)

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 and developing policy in the Aboriginal and Torres Strait Islander health sector by making research and other knowledge readily accessible (Australian Indigenous HealthInfoNet 2020).
- 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.

- Ninti One, an Aboriginal-owned research and consulting organisation, operates the Remote Australia Online database for research relating to life, economic development and social inclusion in remote parts of Australia (Ninti One Limited 2020). This includes the research findings of the Cooperative Research Centre for Remote Economic Participation, a collaborative research program that ran from 2010 to 2017 (Ninti One Limited 2017).
- The Centre of Best Practice in Aboriginal and Torres Strait Islander Suicide Prevention (CBPATSISP) Clearing House is a repository for programs, services, guidelines, research and resources (including for evaluation) that are considered to be best practice in suicide prevention for Aboriginal and Torres Strait Islander people (CBPATSISP 2020).
- The Australian Institute of Health and Welfare (AIHW) collects and publishes a wide range of information on the wellbeing of Aboriginal and Torres Strait Islander people.
 - It is managing a new Indigenous Mental Health and Suicide Prevention Clearinghouse (and is currently establishing a publicly-accessible research and evaluation register that will hold relevant government-commissioned research and evaluation reports).
 - The Regional Insights for Indigenous Communities website (currently in development and to be launched later in 2020) will provide regional (community-level) data on a range of topics including language and culture, health and use of health services.
 - A new standalone website for reporting against the Aboriginal and Torres Strait Islander Health Performance Framework (developed by the Australian Health Minister's Advisory Council for measuring implementation of the National Aboriginal and Torres Strait Islander Health Plan), also to be launched later in 2020, will provide information on progress against 68 metrics (AIHW, sub. DR123, p. 4).

The Australian Institute of Family Studies (AIFS) previously operated 'Knowledge Circle' — a 'what works' portal that provided research resources on the care, protection and wellbeing of Aboriginal and Torres Strait Islander children, with various resources (including evaluation tools) tailored to be useful to practitioners/service providers, researchers, or policy makers. The Knowledge Circle project was discontinued in 2016, and consequently new research is no longer added to its collection (AIFS 2020).

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) also launched the first stages of an Indigenous Research Exchange in 2020. This Exchange aims to provide a portal for research by Aboriginal and Torres Strait Islander academics and organisations; translate both new and existing research into forms that are accessible, useful and practical, especially for Aboriginal and Torres Strait Islander users; identify and prioritise gaps in the evidence base; and provide grants for research priorities to fill those gaps (AIATSIS 2020c; 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 2020c, 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 governments 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, and other research users, 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 research 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, and incorporating new evaluation evidence into the existing knowledge base — collectively referred to as 'knowledge translation' (box 7.5) — will be critical for improving the use of evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people.

Box 7.5 **What is knowledge translation and what does it mean 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

<i>Stage</i>	<i>Strategy</i>	<i>Examples</i>
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 AHCS, 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 (Accountable Income Management Network, 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 need for a central evidence intermediary

Given the extent to which many domains of social and health policy are interlinked, 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 — an ‘evidence intermediary’ — 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) and work with Aboriginal and Torres Strait Islander organisations to facilitate the application of new lessons at the community and service delivery level (Centre for Evidence and Implementation, sub. DR143, p. 2).

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, an Indigenous Evaluation Clearinghouse for evaluations conducted under the Strategy could be established within an existing independent Australian Government agency.

In the draft background paper, the Commission put forward a number of Australian Government agencies that could take on this role and sought feedback on which agency would be best placed to house an Indigenous Evaluation Clearinghouse.

- AIFS and the AIHW both have experience operating the Closing the Gap Clearinghouse, and synthesising research related to Aboriginal and Torres Strait Islander people’s health and wellbeing.
 - The AIHW currently has the largest public repository of research on improving outcomes for Aboriginal and Torres Strait Islander people. The AIHW commented that ‘the complementary nature of [the AIHW’s existing resources] with the Indigenous Evaluation Clearinghouse could enable a more comprehensive drawing together of evaluation evidence with monitoring and other data, at regional levels where possible, to highlight areas of need and priorities for future investments’ (sub. DR123, p. 2).
 - AIFS has also demonstrated skills in making sense of evidence and knowledge translation (for example, its *Child Family Community Australia* information exchange and its *Knowledge Circle* information portal).

-
- AIATSIS has an exclusive focus on research involving Aboriginal and Torres Strait Islander people (and strong relationships with Aboriginal and Torres Strait Islander organisations). It is responsible for the *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research* and has recently launched an Indigenous Research Exchange, which will include a knowledge translation function. AIATSIS is also governed by a majority-Indigenous Council. The National Centre of Indigenous Excellence (sub. DR158, p. 22) and Royal Australian and New Zealand College of Psychiatrists (sub. DR138, p. 4) highlighted the Council as a key reason for AIATSIS to have responsibility for the Clearinghouse.
 - The Productivity Commission could also take on the role. It has experience in synthesising research and data, and in working with Aboriginal and Torres Strait Islander people and organisations (such as through the *Overcoming Indigenous Disadvantage* report series). The Alliance of First Nations Independent Education and Training Providers (sub. DR141, p. 6) and Yulang Indigenous Evaluation (sub. DR150, p. 20), said the Commission's independence and the broad scope of its work meant it was the most appropriate agency to house the Clearinghouse. There are also some synergies with the Commission's new responsibilities under the National Agreement on Closing the Gap (the National Agreement) — the Commission will be responsible for developing and maintaining a Dashboard of data and preparing an annual data compilation report (underpinned by an Information Repository) to inform reporting on progress against the priority reforms, targets, indicators and trajectories. The Commission will also undertake a triennial review on progress towards Closing the Gap (which will draw on evaluation and other relevant evidence) (JCOCTG 2020, pp. 40–41).

Some participants suggested that an Aboriginal and Torres Strait Islander community controlled organisation should be responsible for hosting and operating the clearinghouse:

... we recommend that the proposed Indigenous Evaluation Clearinghouse be housed by a strongly-governed First Nations organisation. Such an approach aligns with and will enhance 'data sovereignty' ... It also aligns with the new National Agreement on Closing the Gap's outcomes, particularly building the community-controlled sector and Aboriginal and Torres Strait Islander-led data. A First Nations organisation is better placed to understand the cultural requirements for this important role, and ensure data is shared in a way that is relevant for First Nations people. (Community First Development, sub. DR134, p. 9)

... an independent Aboriginal and Torres Strait Islander community-controlled evidence intermediary [is recommended because of] the power that controlling data, evidence and evaluation gives to governments. For the evidence intermediary to be effective, it must shift power from governments to communities and ACCOs, working towards data sovereignty and capacity-building for self-determined evaluation. ... For these reasons, we envisaged that the independent central agency and evidence intermediary would be independently run and self-determined — rather than a government agency (i.e. proposed OIPE) advised by a Council. (National Aboriginal and Torres Strait Islander Legal Services, sub. DR160, p. 5)

While a community controlled Indigenous Evaluation Clearinghouse would align with Priority Reforms Two (building the community controlled sector) and Four (shared access to data and information at a regional level) of the National Agreement, the Commission is

of the view that establishing the Clearinghouse in an independent statutory authority is important for stability and longevity (as noted in box 7.6, this is critical for the effective operation of an evidence intermediary). Another way to support Priority Reform Two of the National Agreement would be for the host agency to contract out some of the translation, synthesis and meta-analysis work to Aboriginal and Torres Strait Islander researchers, including community controlled organisations.

Based on the ‘common ingredients’ for successful evidence centres (box 7.6), no one agency stands out as being best placed to house an Indigenous Evaluation Clearinghouse. However, there would be benefits to having a clear line of sight from the Indigenous Evaluation Clearinghouse to the Productivity Commission’s new data repository and reporting functions under the National Agreement (so that synthesised evaluation evidence could be easily fed into the information repository and triennial review). It will also be important to minimise duplication of existing information repository and knowledge translation functions.

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 the 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 (as proposed in the Strategy, action 8) is in line with the National Health and Medical Research Council’s (NHMRC) suggestion in *Keeping research on track II* — 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).

Box 7.6 The ‘common ingredients’ of successful evidence centres

The Nesta Foundation (Nesta), formerly the National Endowment for Science, Technology and the Arts, is an innovation-focussed research and advocacy charity in the United Kingdom. Nesta hosts the Alliance for Useful Evidence (the Alliance) — a network of individuals (from government, universities, charities, businesses, and local authorities in the UK and internationally) that champions the use of evidence in social policy and practice. Nesta and the Alliance have assisted with the development of several independent evidence centres, and identified the following ‘common ingredients’ as being important to the establishment of a successful evidence centre.

1. **Clear objectives:** Knowledge of the centre’s intended user groups and the context they are working within. Understanding of the relationship between evidence producers and users, and of the key drivers that influence users’ decision-making. Clear, observable and measureable outcomes to work towards, and an evidence-informed theory of change.
2. **Robust organisational development:** Commitment to create an independent and sustainable organisation with effective governance and the right mix of skills and experience. Stability and longevity of funding, so the centre can operate over a timescale that will be sufficient to make a difference. Engagement with intended users to inform decisions about organisational design.
3. **Engaged users:** Understanding of users’ evidence needs (what evidence is in demand; which ‘knowledge mobilisation’ techniques will help the most?). Collaborative work systems that increase users’ capability, motivation and opportunity to employ evidence in decision-making.
4. **Rigorous curation and creation of evidence:** A robust and transparent approach to selecting and generating high-quality evidence for users. This can include: defining what the centre considers to be ‘evidence’ and what evidence the centre will curate; applying standards of evidence and mapping existing evidence; synthesising the most useful evidence; facilitating primary research to fill evidence gaps; building sector-wide capacity to create robust evidence.
5. **A focus on impact:** Commitment to learn from successes and failures, so that through continuous quality improvement the centre can increase its effectiveness in achieving objectives.

Source: Bazalgette (2020).

Several participants highlighted the importance of reciprocity in Aboriginal and Torres Strait Islander relationships (chapter 1) and delivering research results back to Aboriginal and Torres Strait Islander communities (including in local languages where appropriate) to provide accountability and share the benefits of research.⁴⁰ Sharing research results with participants is also an important aspect of ethical research and accountability more broadly (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 pointed out, publication of reports is not always the only appropriate method of communicating evaluation findings (including to Aboriginal and Torres Strait Islander communities). Other methods 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

⁴⁰ For example: Aboriginal Health Council of Western Australia (sub. 42, p. 3); Accountable Income Management Network (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).

organisations early in the evaluation planning process about how results are best shared, including for implementation in the community where practicable, 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 2016a, p. 10) that debriefs department staff on key learnings from an evaluation, including comments about the strengths and challenges of the evaluation process itself.⁴¹ 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 proposes that all evaluations under the Strategy should include a clear and concise summary of the evaluation report that documents how the planning, commissioning and conduct of the evaluation adhered to the Strategy’s principles (Strategy, action 8). Agencies’ management responses (which could be included as part of the evaluation report or published separately, but in any case should be published shortly after the completion of the evaluation) should include an explanation about 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 (Strategy, action 9).

Ensuring that evaluation findings feed into decision-making processes

For the Strategy to achieve its objective, the increased accessibility and usefulness of evaluation findings also needs to be complemented by agencies having stronger incentives to use those results when developing and implementing future policies and programs. A number of participants expressed concerns that higher-quality evaluations, conducted in accordance with the Strategy, would be put on the shelf to ‘gather dust’ (rather than being used to achieve meaningful policy change).

While the widely accepted intention of policy and program evaluation is to feed back into implementation and future design, this has not always been the case for Australian government programs which impact upon or target Indigenous peoples. (Accountable Income Management Network, sub. 15, p. 12)

Evaluation must be a living process, not one that produces unused evaluation products that sit on the shelves of policy makers. (Empowered Communities, sub. 41, p. 6)

⁴¹ The former Department of Industry, Innovation and Science (now the Department of Industry, Science, Energy and Resources) also has this as part of its *Evaluation Strategy 2017–2021* (DIIS 2017, p. 33).

It is ... imperative that evaluation findings are translated into practice and embedded in the future decision-making cycle for subsequent policy design through iterative cycles. (National Aboriginal and Torres Strait Islander Health Standing Committee, sub. 104, p. 4)

To ensure the objective of the Strategy is achieved, governments must genuinely move toward focusing on policy outcomes and effectiveness, and view evaluation as a means of continuous quality improvement. Without this, the cycle of producing reports and strategies that result in no real change will continue ... (Australian Medical Association, sub. DR116, p. 2)

... the strategy and guide are all very well, but unless government agencies are held accountable for their commitment to evaluation and the findings that arise, then change will be limited. (Ninti One Limited, sub. DR118, p. 2)

Several participants raised the example of the Indigenous Legal Assistance Program (ILAP), which was reviewed during its 2015–2020 funding cycle. The review recommended that community controlled Aboriginal and Torres Strait Islander legal services continue to be funded through standalone, specific funding mechanisms to promote self-determination and deliver culturally appropriate legal services (Cox Inall Ridgeway 2019, p. 16). However, the 2019-20 Budget stated that ILAP funding would be mainstreamed into a single mechanism for State and Territory legal assistance funding (Commonwealth of Australia 2019, p. 55), through the *National Legal Assistance Partnership 2020–2025* (though the Partnership Agreement states that funding for ATSILS will be ‘quarantined’: COAG 2020, pp. 16–17). National Aboriginal and Torres Strait Islander Legal Services argued that this example ‘demonstrates the lack of influence that evaluations can have on government decision-making and points to the importance of the Indigenous Evaluation Strategy including government accountability mechanisms’ (sub. 97, p. 17).

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 2020c, pp. 123–124). The Independent Review of the APS commented that this is only done in an ad hoc fashion in the Australian Government, and suggested that a more systematic approach was needed (DPMC 2019e, p. 270).

Some Australian jurisdictions do have evaluation built explicitly into the budget cycle (chapter 2, table 2.1). For example:

- In New South Wales, all program-related submissions to Cabinet seeking funds need to include details of: when the program was last evaluated (if the submission is to renew 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).
- Western Australia has a ‘sunset clause’ system for all programs with an annual cost of at least \$5 million, meaning 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 (Department of Treasury (WA) 2015, p. 5).

-
- The Northern Territory Government also has a sunset clause, which applies to programs worth \$1 million or more in any one year and sets funding for no more than five years, with ongoing funding to be informed by evaluation (Department of Treasury and Finance (NT) 2020).

There are no similar requirements around the formal use of evaluation results in policy making for the APS. However, some Australian Government agencies set out, in their evaluation strategies, an expectation that evaluation results will feed into future policy making. For example:

- the former Department of Industry Innovation and Science (now the Department of Industry, Science, Energy and Resources) *Evaluation Strategy 2017–2021* states that evaluation results should ‘inform decision-making and policy development’ (2017, p. 9), and that policy/program executives hold responsibility for implementing evaluation recommendations (2017, p. 31)
- the former Department of Infrastructure and Regional Development (now the Department of Infrastructure, Transport, Regional Development and Communications) *Evaluation Strategy 2016–2021* notes that ‘evaluation-related products … being used to inform decision-making’ is a necessary metric of that strategy’s success (2016b, p. 11).

Formal mechanisms or requirements to substantively implement the results of evaluations are rare. The clearest example of an agency having a relevant policy was the Office for Development Effectiveness — it previously published on the DFAT website an annual *Review of Uptake of ODE Recommendations*, which tracked the progress of initiatives agreed in management responses to the recommendations of the ODE’s strategic evaluations.

A number of participants suggested that the Strategy should include accountability mechanisms for the use of evaluations, such as a system for tracking the use of evaluation results, or formal requirements for evaluation findings to be incorporated into future policy and program design:

… 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 … (NSW Aboriginal Land Council, sub. 91, p. 2)

… it is essential that the Strategy should be expanded to consider … how those responsible for managing evaluations can best manoeuvre through the ‘political’ process of ensuring the take-up of evaluation outcomes. This is likely to include clear governance structures for each evaluation, which include the representation of Indigenous people, able to hold policymakers and program implementers accountable for reasonable utilisation of evidence. (Institute for Human Security and Social Change, sub. DR119, p. 2)

All evaluation recommendations should be mandated as KPIs for Ministers and Secretaries of the relevant department, program or policy. We must ensure that leaders within the public service are accountable for evaluation results and enactment of evaluation recommendations, and continue along the policy cycle into policy and program improvement. … If the recommendation cannot be

met, sufficient evidence must be provided to demonstrate why this has not occurred and what action plan is in place to rectify the issue. (VACCHO, sub. 44, pp. 22–23)

Given the realities of the policy-making process, it is not possible in practice to guarantee that evaluation results, or any other type of evidence, will have a particular *effect* on the design of any given policy or program — even where there is a requirement to consider evidence during the policy design stage (Tune 2010). And some evaluations will not be well-suited to generating findings that can be easily actioned.

However, requiring agencies to publicly respond to evaluation findings (including a plan for incorporating those conclusions into future policy or program design) will increase agencies' accountability to Aboriginal and Torres Strait Islander people, organisations and communities, and the Australian community overall. It will also make it easier for policy makers and other evaluation users to track the changes made as a result of evaluations.

Strategy actions for better use of evaluation evidence

The Strategy has four actions designed to improve the accessibility, usability and transparency of evaluation findings.

- 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 (and in this case a summary of the evaluation's findings should be published, including an explanation for why the full evaluation report could not be published) (Strategy, action 7).
- Evaluation reports under the Strategy should include a short plain English summary. Other ways of sharing evaluation findings should be agreed with the relevant Aboriginal and Torres Strait Islander people, organisations and communities at the evaluation planning stage. Evaluation reports (or summaries of findings) should also document how the planning, commissioning and conduct of the evaluation adhered to the principles of the Strategy (Strategy, action 8).
- Australian Government agencies should publish a management response to each evaluation or summary report (Strategy, action 9). 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.
- The Australian Government should establish an online clearinghouse for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people (Strategy, action 10). The clearinghouse should sit within an existing independent statutory authority, and be responsible for coordinating the synthesis of new evaluation evidence into the existing evidence base, and for translating the knowledge into accessible forms. All published evaluations and summary reports should be made available to the clearinghouse (action 7).

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 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 — has meant greater external scrutiny. However, there are no formal requirements for evaluation under the Framework (chapter 2), and some participants said 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.

Under current arrangements, the only external scrutiny of Australian Government agencies' evaluation processes is that undertaken by the Australian National Audit Office (which recently released a report on performance reporting and evaluation under the Indigenous Advancement Strategy Framework)⁴² and, at times, by the Commission (which 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)).

External oversight to change incentives

One way to strengthen accountability is to have external oversight. Keeping tabs on what agencies are evaluating, how they are conducting evaluations, and their use of evaluation findings in policy and program development can encourage agencies to conduct evaluations that are useful and high quality, and to make greater use of those evaluations (chapter 2; appendix D).

The Australian Government has already acknowledged the need for the Strategy to have independent external oversight by giving the Commission the role of monitoring the

⁴² 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).

performance of Australian Government agencies against the Strategy. The Letter of Direction for this project 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’. 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).

The Commission, as a statutory body, has the independence from the day-to-day activity of government necessary to objectively assess agencies’ performance against the Strategy. However, because the Commission does not have powers to direct government agencies in the exercise of their day-to-day functions, it could not compel agencies to implement the Strategy. The key method of incentivising agencies to lift the bar on the quality and usefulness of evaluations would be transparent public reporting on agencies’ performance against the Strategy’s principles and actions.

The Commission’s proposed approach to monitoring agencies’ performance, which includes biennial public reporting and a progression pathway for assessing agencies’ performance, is set out in chapter 10.

An Indigenous governance mechanism

To align with the Strategy’s overarching principle of centring Aboriginal and Torres Strait Islander people — and the Australian Government’s commitment to work in partnership with Aboriginal and Torres Strait Islander peoples on matters affecting their lives and communities — oversight arrangements (and those for setting evaluation priorities, chapter 6) require Indigenous governance to bring an Indigenous lens. An Indigenous lens will bring lived experience and expertise in the development, delivery and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.

A number of participants commented on the importance of having an Aboriginal and Torres Strait Islander governance mechanism for the Strategy. For example:

... we strongly support the implementation of an Aboriginal governance structure [being] included as a main component of the overall strategy. (Aboriginal Health Council of South Australia, sub. 61, p. 2)

... the role of First Nations Peoples [should] 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. (Indigenous Community Volunteers, sub. 88, p. 8)

Oversight of the Strategy should also include governance arrangements that ensure effective engagement with Aboriginal and Torres Strait Islander people more generally and on an ongoing basis. (National Voice for our Children (SNAICC), sub. DR163, p. 8)

Some of the participants at the ‘Culture, capability and governance’ roundtable (hosted by the Commission in November 2019) were of the view that the Strategy’s governance should include a partnership arrangement with Aboriginal and Torres Strait Islander people. Several

participants commented that advisory or other consultative bodies often have little influence,⁴³ hence the importance of formal partnership or co-design arrangements.

There are a number of examples of Aboriginal and Torres Strait Islander-led advisory, oversight and representative mechanisms in the Australian, state and territory governments (box 7.7).⁴⁴ Some Australian Government agencies have Indigenous-majority Boards or Councils specified in their legislation, including AIATSIS, Indigenous Business Australia, and the Indigenous Land and Sea Corporation. The Victorian Government is currently developing an Aboriginal-led evaluation and review mechanism to track the public sector's progress against the *Victorian Aboriginal Affairs Framework 2018–2023* (DPC (Vic) 2019). And the WA Government has proposed establishing a statutory office for accountability and advocacy in Aboriginal affairs, which may include a monitoring and evaluation function (DPC (WA) 2018).

It is important that any governance arrangements put in place under the Strategy are not tokenistic, but reflect the principles of shared decision making (as set out in Priority Reform One of the National Agreement). This will be important for centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges and accountability to Aboriginal and Torres Strait Islander people. The unit within the Commission responsible for overseeing the Strategy would therefore need to partner with Aboriginal and Torres Strait Islander people to determine government-wide evaluation priorities and to monitor the performance of agencies against the Strategy.

Other roles will be important if the Strategy is to be effective

It is the Commission's view that if the focus is only on monitoring agencies' performance — without support for agencies to improve their evaluation cultures and practices, and their processes for engaging with Aboriginal and Torres Strait Islander people — there is a real risk that agencies will view the Strategy as little more than a compliance hurdle.

It will be important to have evaluation leadership (an 'evaluation champion') to build a culture that values evidence, builds evaluative capability, shares knowledge (including about what has not worked), and learns from evaluation findings.⁴⁵ As Alexander and Thodey (2018, p. 13) said, 'the tone is set at the top'. And in a survey of Australian evaluators by

⁴³ For participants' comments on advisory or other consultative bodies not having substantive influence, see: Tony Dreise et al., (sub. 33, p. 3); Lowitja Institute (sub. 50, p. 8); Royal Australian College of Physicians (sub. 77, p. 4); Professor Maggie Walter (sub. 112, pp. 3–4).

⁴⁴ An international example of an Indigenous oversight mechanism is the Independent Māori Statutory Board. Under the *Local Government (Auckland Council) Amendment Act 2010* (NZ), the Board has a statutory responsibility for promoting 'Issues of Significance' to Māori in Tāmaki Makaurau (Auckland) and monitoring the Auckland Council's performance in responding to those issues (appendix D, box D.8).

⁴⁵ Don Weatherburn also noted that 'Commitment to rigorous and objective policy and program evaluation starts at the top. If the CEO of an organisation and his or her senior offices are not committed [to] rigorous evaluation and its corollary (preparedness to accept disappointing results), then no amount of commitment at lower levels of the organisation will make it happen' (sub. DR114, p. 3).

Maloney (2017, p. 30), most respondents indicated that leadership had more of an impact on evaluation use than agency evaluation policies and processes.

Box 7.7 Aboriginal and Torres Strait Islander governance/advisory functions in Australian jurisdictions

Aboriginal and Torres Strait Islander Elected Body (ATSIEB) (ACT)

Established under the *Aboriginal and Torres Strait Islander Elected Body Act 2008* (ACT), the ATSIEB is made up of seven Aboriginal and Torres Strait Islander elected members (with all Aboriginal and Torres Strait Islander ACT residents eligible to vote for membership). Its functions include monitoring and reporting on the effectiveness and accessibility of programs and services conducted by government agencies in the ACT, and providing advice to government agencies on the appropriate representation of Aboriginal and Torres Strait Islander people through government consultative bodies. The ATSIEB holds Senate Estimates-style public hearings and reports on these to the Minister for Aboriginal and Torres Strait Islander Affairs, with the Minister required to present those reports to the Legislative Assembly and formally respond to the ATSIEB.

Western Australian Aboriginal Advisory Council (WAAAC)

The WAAAC is a statutory board (established under the *Aboriginal Affairs Planning Authority Act 1972* (WA)), currently comprising 11 members who were nominated by Aboriginal community leaders from across the State. The WAAAC's role is to advise the Department of the Premier and Cabinet (DPC) and the Minister for Aboriginal Affairs on matters relating to the interests and wellbeing of Aboriginal people in WA (DPC (WA) 2020). The DPC and the Minister must consider the WAAAC's views and recommendations, but are not bound to follow these. The WAAAC also has a direct relationship with the Aboriginal Affairs Co-ordinating Committee.

National Indigenous Australians Agency (NIAA) Indigenous Evaluation Committee

The Indigenous Evaluation Committee (IEC) is made up of 3–5 external members (appointed by the NIAA's CEO), and the NIAA's Deputy CEO. Currently, two of the four external members are Aboriginal (though this is not a formal requirement). The Committee's objective is to 'strengthen the quality, credibility, and influence of the evaluations of policies and programs led by the NIAA through the provision of independent strategic and technical advice'. Responsibilities include: overseeing implementation of the Indigenous Advancement Strategy Evaluation Framework; advising the NIAA on the preparation of its forward evaluation program; providing technical advice, guidance and quality assurance on evaluations; and reviewing and summarising cross-cutting evaluation reports (NIAA 2019b). The IEC reports to the NIAA's CEO.

Agencies will also need guidance on how to implement the Strategy, and to be incentivised to evaluate with the purpose of improving policy and program outcomes for Aboriginal and Torres Strait Islander people. This means acknowledging experimentation and innovation, 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 (with that information being used to reform policies and ultimately to improve lives).

Sharing the learnings from evaluation — and encouraging their integration into agencies' practices — will also be important for creating a culture where evaluation results are valued, and for strengthening agencies' capability to effectively evaluate policies and programs affecting Aboriginal and Torres Strait Islander people. Evaluation learnings can also be used outside the APS, including by: Aboriginal and Torres Strait Islander people, organisations and communities; state and territory governments; independent evaluators; and academics.

Many participants agreed that an independent central evaluation body was required for the Strategy to be effective — not only to hold agencies accountable for planning, conducting and publishing high-quality evaluations, but also to provide practical guidance on evaluating in line with the Strategy, help to build evaluative and cultural capability across the APS, and to identify evaluation priorities (box 7.8).

In the draft background paper and Strategy, we proposed that the Australian Government establish an Office of Indigenous Policy Evaluation (OIPE), in an independent statutory agency, with the following functions:

- have stewardship of the Strategy
- champion the value of evaluation across the APS
- advise Australian Government agencies on what the Strategy requires in practice⁴⁶
- identify evaluation priorities for policies and programs affecting Aboriginal and Torres Strait Islander people (chapter 6)
- identify potential areas for evaluations or systematic reviews on cross-cutting topics (those that involve the work of more than one agency)
- oversee agencies' Three Year Evaluation Forward Work Plans (chapter 6)
- coordinate with the Head of Evaluation Profession on capability-building opportunities⁴⁷ (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 was proposed on the basis that ideally (for reasons of efficiency and effectiveness) all the functions to support the Strategy, and ensure it is effective, would be located in one place. This also means that an Indigenous governance mechanism could work with the OIPE across all of its functions. The central clearinghouse and knowledge translation function could also (but would not have to) be a responsibility of the OIPE.

⁴⁶ A number of participants supported an office providing guidance to agencies on conducting evaluation in line with the principles of the Strategy — Ninti One Limited (sub. DR118, p. 2); Ernst & Young (sub. DR140, p. 5); National Voice for our Children (SNAICC) (sub. DR163, p. 12); Department of Health (sub. DR176, p. 4).

⁴⁷ It is not yet clear when a Head of Evaluation Profession will be appointed. However, in the interim an OIPE could support the creation and coordination of evaluation and cultural capability-building opportunities.

Box 7.8 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:

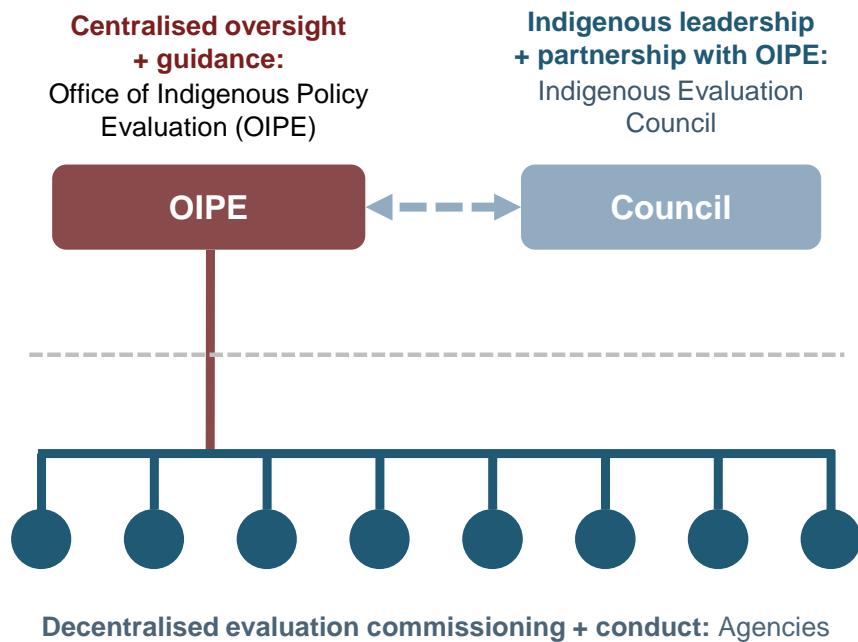
- support and develop evaluation and quality improvement literacy;
- ensure a consistent approach to evaluation across departments;
- monitor and evaluate practice; and
- translate evaluation outcomes more broadly into policy and practice. (sub. 50, p. 10)

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)

The draft background paper and Strategy also proposed an Indigenous Evaluation Council (Council) (with the majority of members being Aboriginal and Torres Strait Islander people), envisaging that the Council would formally partner with the OIPE on its monitoring, reporting and priority-setting functions, and provide strategic guidance on the OIPE's other functions (figure 7.1).

Figure 7.1 Governance arrangements for the Strategy



The feedback on the draft background paper and Strategy (box 7.9) was that some participants supported the OIPE and Indigenous Evaluation Council, with a number arguing that the National Agreement should be the basis for the formation, roles and operation of the two bodies (with a focus on partnership and shared decision making with Aboriginal and Torres Strait Islander people, and transforming mainstream institutions). Some participants also said that, to accord with the National Agreement's principles around leadership, an OIPE should be staffed and/or led by a majority of Aboriginal and Torres Strait Islander people:⁴⁸

[SNAICC recommends establishing] minimum Aboriginal and Torres Strait Islander employment targets for staffing the OIPE, including ensuring that its Chief Executive or leader is an Aboriginal and/or Torres Strait Islander person. (National Voice for Our Children (SNAICC), sub. DR163, p. 5)

The QFCC believes the proposed Office of Indigenous Policy Evaluation ... should be led and staffed — as well as governed — by a majority of Aboriginal and Torres Strait Islander people. This will ensure Aboriginal and Torres Strait Islander people fully participate in evaluation, rather than be ‘engaged with’ ... (Queensland Family and Child Commission, sub. DR136, p. 6)

⁴⁸ See also: Alliance of First Nations Independent Education and Training Providers (sub. DR141, p. 6); Beyond Blue (sub. DR121, p. 1); Wuchopperen Health Service Limited (sub. DR117, p. 1); Yulang Indigenous Evaluation (sub. DR150, p. 20).

Box 7.9 Comments on the proposed governance arrangements

Beyond Blue:

The inclusion of an Indigenous-led evaluation champion to oversee the Strategy and Indigenous-led Evaluation Council to govern the Strategy is a major strength. (sub. DR121, p. 1)

ANTaR:

ANTaR strongly endorses the proposal for an Office of Indigenous Policy Evaluation (OIPE) and Indigenous Evaluation Council. ... There has been criticism of reports like the Indigenous Expenditure Report where significant expenditure is discussed without a nuanced analysis or appreciation of the value, or otherwise, being delivered by that expenditure. The OIPE could be the body that provides this nuanced analysis with its particular expertise on Aboriginal and Torres Strait Islander policy, along with the maintenance of a trusted network of relationships with First Nations organisations. In turn, an Indigenous Evaluation Council can ensure that evaluation practices are grounded in Aboriginal and Torres Strait Islander values and can help to focus these practices within the intercultural context. (sub. DR124, pp. 7–8)

National Voice for our Children (SNAICC):

SNAICC is supportive of the establishment of the proposed Office of Indigenous Policy Evaluation (OIPE). ... SNAICC acknowledges plans for the office to sit within an existing independent statutory authority in its establishment phase however would be supportive of the office being a standalone independent authority into the future. SNAICC is supportive of the OIPE promoting a culture of accountability and requiring agencies to demonstrate their active efforts in adopting the key principles and actions within the Strategy. (sub. DR163, p. 8)

Riverina Murray Regional Alliance:

The proposed Office of Indigenous Policy Evaluation (OIPE) is positioned as an “evaluation champion” for the Strategy, however, has only advisory, and no decision-making power. Granting the OIPE decision-making power would allow it to hold the government more closely to account. (sub. DR149, p. 15)

Lowitja Institute:

The recommendation to create the Office for Indigenous Policy Evaluation (OIPE) is likely to be controversial because it potentially creates another government body with little power and authority over other agencies to enforce the principles and practices outlined in the Strategy. Our view is that an Aboriginal and Torres Strait Islander governed body to drive the Strategy is important ... we support establishing strong Aboriginal and Torres Strait Islander governance and advisory mechanisms to support the delivery of the Strategy. (sub. DR157, p. 4)

The National Health Leadership Forum:

The NHLF supports the proposal for an overseeing role of the IES across Commonwealth agencies, however it is not fixed on the [OIPE/Council] structure, but rather whatever the structural form, it must be Indigenous led. (sub. DR146, p. 3)

NSW Aboriginal Land Council:

We are concerned the proposal [for an OIPE and Council] may create additional bureaucracy, divert resources away from the community-controlled sector, and potentially duplicate other work. We suggest that consideration is given to effective mechanisms to ensure genuine partnership with Aboriginal and Torres Strait Islander people in the design and implementation of any evaluation strategy ... the partnerships arrangements of the National Agreement should be reflected in any governance arrangements related to evaluation. (sub. DR122, p. 3)

National Aboriginal Community Controlled Health Organisation:

[The proposed governance arrangements] have the potential to reduce the accountability of Government Agencies, as well as that of the Aboriginal and Torres Strait Islander organisations they fund, and increase workloads including duplication. ... The proposed [OIPE] seems to replicate the failed Office of Indigenous Policy Coordination (OIPC), which sat within the Department of Immigration and Multicultural and Indigenous Affairs and was formally disbanded in August 2011. (sub. DR170, p. 3)

However, some participants did not support the proposed governance arrangements, arguing that setting up additional government bodies would add unnecessary complexity and administrative burden to the work of service providers, and would take away resources and responsibility from the community controlled sector (box 7.9). And others were agnostic about specific governance mechanisms, but stressed that any structural reform to support the Strategy’s implementation must be Indigenous-led, and must respect existing partnership arrangements under the National Agreement.

Resourcing concerns are not surprising, particularly given the current budgetary environment. However, the Commission is of the view that an OIPE and Indigenous Evaluation Council would be key components for the success of the Strategy — not only for monitoring agencies’ progress but also in terms of supporting good evaluation practice within the APS (catalysing growth in evaluative culture), better disseminating the learnings from evaluation, and ensuring that evaluations are directed at the areas in most need of evidence. The Commission envisages the OIPE to be a small office (requiring between 4–5 full-time equivalent staff), and the costs (including those for Council members) to be significantly outweighed by the benefits of this architecture.

Where should the OIPE be located?

On the question of where the OIPE might be located, several participants supported the Productivity Commission taking on the additional functions to support the Strategy (box 7.10).

Some of the reasons identified for the Commission taking on the role of the OIPE included: legislative independence; influence across the APS; and longevity for the Strategy (given that the Commission is underpinned by legislation). Others were that:

- the Strategy’s monitoring and public reporting role already sits with the Commission
- there are synergies between the work of the OIPE and the Commission’s work under the National Agreement (as noted above, the Commission will report every three years on progress against the Priority Reforms and targets, including drawing on evaluation and other evidence, and will maintain an information repository)
- the Commission has an appointed Indigenous Commissioner, who could potentially sit on the Indigenous Evaluation Council
- the Commission has well-established, transparent engagement processes in place. It is also a champion for evidence-based policy.

Box 7.10 Participants' views on the location of an Office of Indigenous Policy Evaluation and Indigenous Evaluation Council

Empowered Communities:

If an OIPE is established with functions as proposed it will be crucial that it sits within the broader Productivity Commission. This would both provide a productivity link and better enable the power and influence that it will require to do its job effectively. It will need legislative 'teeth' to harness all relevant parties, governments, funded service providers and others, and ensure they follow through on commitments under the Strategy. Establishment in another, even relatively arm's length, government agency, without its own legislative base and direct productivity link would lessen its strength to act as it will need to, to support Indigenous people and government to close the gap. (sub. DR171, p. 5)

Independent Members of the NIAA Indigenous Evaluation Committee:

Finding a position for the [Council] that is both independent and influential is critically important, but getting this balance right also has implications for the longevity of the IES and its infrastructure. These governance arrangements, and the IES principles and behaviours to be implemented, need to be sustained over the longer term. We therefore support the proposal that the [OIPE] and the Council be located in an independent but also influential agency such as the Productivity Commission. We further suggest that members of the Council be appointed through a collaborative process involving the [Coalition of Peaks], Productivity Commission and the National Indigenous Australians Agency. (sub. DR115, pp. 2–3)

Royal Australian and New Zealand College Of Psychiatrists:

The RANZCP recommend that the Office of Indigenous Policy Evaluation be situated within the National Indigenous Australians Agency (NIAA). Since its establishment the NIAA has taken on key government priority areas for policy and reform, including the Aboriginal and Torres Strait Islander Mental Health and Suicide Prevention Advisory Group, National NAIDOC Committee and been closely involved in the renewed Closing the Gap strategy. As evaluations occur across government the NIAA could provide a point of coordination for conceptualisation of evaluation in agencies such as AIHW, through to publication through agencies such as AIATSIS. (sub. DR138, p. 5)

Alliance of First Nations Independent Education and Training Providers:

The Alliance believes it is important that the new central advisory and oversight function under the Strategy — an [OIPE] — not be inadvertently limited in its scope [therefore should be] placed within an agency that is independent of the evaluation process and is not focused on a narrow range of policy issues. For these reasons, the Alliance believes the best agency to house the OIPE is the Productivity Commission. The Productivity Commission could obtain the required Indigenous staff to support this role ... (sub. DR141, p. 6)

Yulang Indigenous Evaluation:

An Office of Indigenous Policy Evaluation location in the Productivity Commission is preferable to any other currently existing agency. It is established, has its focus best use of resources, has established processes and profile for reporting on progress and has staff skilled in evaluation. However, more staff, including Aboriginal and Torres Strait Islander staff and those highly training in working with Aboriginal and Torres Strait Islander people toward self-determination must be employed. (sub. DR150, p. 20)

National Centre of Indigenous Excellence:

We support [the establishment of an Indigenous Evaluation Clearinghouse] and believe AIATSIS is [the appropriate agency] ... In order to best position the OIPE to help drive organisational change, this function should be incorporated within the OIPE, or at least responsibility for the coordination of this activity should be vested with them ... (sub. DR158, p. 22)

Others suggested the OIPE would be more appropriately housed within the NIAA (Royal Australian and New Zealand College of Psychiatrists, sub. DR138, p. 5) or AIATSIS

(National Centre of Indigenous Excellence, sub. DR158, p. 22). This likely reflects their more established cultural capability and leadership working with Aboriginal and Torres Strait Islander people.

The NIAA currently has, as one of its statutory functions, analysing and monitoring the effectiveness of programs and services for Aboriginal and Torres Strait Islander people, including those delivered by other agencies.⁴⁹ It also has a relatively sophisticated internal evaluation unit. However, on monitoring, the NIAA may be seen as ‘marking its own homework’ (because it designs and delivers policies and programs, and also evaluates them). Many participants highlighted the importance of independence (from policy and program agencies) where oversight of the Strategy is concerned, not only to mitigate conflicts of interest but also to ensure that reporting on agencies’ performance against the Strategy is credible.

AIATSIS is an independent statutory authority, so it would be able to provide oversight at arm’s length from policy and program agencies. It would also bring strong Indigenous leadership (as mentioned above, it is governed by an Indigenous-majority Board) and expertise on ethics and protocols for research with Aboriginal and Torres Strait Islander people. AIATSIS would be well-placed to make sense of evaluation findings, assess the existing evidence base for gaps, and identify priority areas for evaluation. However, it does not have organisational experience in performance monitoring or specialist expertise in evaluation. It would need to broaden its current research remit (and capacity) to ensure it can oversee evaluation across a wider range of policies and programs.

Other participants argued that centralised oversight of the Strategy should be entirely independent of the Australian Government, and should have strong Aboriginal and Torres Strait Islander governance (NATSILS, sub. DR160). A non-government Aboriginal and Torres Strait Islander organisation would bring valuable insights on what centring means for evaluating policies and programs, and would also be well-placed to make sense of evaluations findings (particularly in terms of working with Aboriginal and Torres Strait Islander communities and organisations to implement relevant learnings) and to identify priority areas for evaluation. It would also have the benefits of Indigenous leadership and strong connections with Aboriginal and Torres Strait Islander people, organisations and communities. However, a non-government organisation would not have the stability of a statutory authority and is also likely to have to build expertise in performance monitoring and evaluation of government agencies.

A second-best option is to separate out some of the functions of the OIPE. For example, the evaluation unit within the Department of Finance could take on the role of evaluation champion and provide advice to agencies about applying the Strategy in practice. The NIAA could also take on the capability building role in collaboration with the Australian Public Service Commission. The Productivity Commission would retain the roles specified in the

⁴⁹ Section (e) vii, *Order to Establish the National Indigenous Australians Agency as an Executive Agency 2019* (Cwlth). This means that the NIAA has the ability to commission or conduct evaluations of policies and programs delivered by other Australian Government agencies, whether on its own or in partnership with the agency in question.

Letter of Direction — to monitor agencies’ progress in implementing the Strategy and evaluation priority-setting.

However, as noted earlier, the first best option is to have all the proposed functions to support the Strategy in one place. An OIPE and an Indigenous Evaluation Council working together are considered essential architecture for an effective Strategy.

The OIPE and the Indigenous Evaluation Council — how would it work?

The OIPE and the Council, in partnership, would:

- produce biennial reports on the performance of Australian Government agencies against the Strategy. These reports should: identify good and/or innovative evaluation practice, highlight key lessons learned and remaining gaps in the evidence base, analyse how effective the Strategy has been in encouraging high-quality and useful evaluations, and provide recommendations on how the implementation of the Strategy could be improved. Chapter 10 outlines a proposed approach to assessing agencies’ performance against the Strategy, which includes rolling Three Year Evaluation Forward Work Plans, agencies’ depth of engagement with Aboriginal and Torres Strait Islander people in prioritising, planning, undertaking and sharing evaluations, and the quality of evaluations
- update the whole-of-Australian Government evaluation priorities under the Strategy, including identifying areas for cross-agency evaluations.

On a day-to-day basis, much of the OIPE’s monitoring work (including collecting evaluation documentation and rolling Three Year Evaluation Forward Work Plans, and tracking agencies’ progress against various Strategy requirements) would be administrative in nature and not require the input of the Council. The Council could, however, provide strategic guidance to the OIPE on the direction and focus of:

- the OIPE’s stewardship of the Strategy
- building evaluation and cultural capability
- its advice to agencies on evaluation planning, commissioning, conduct, publication and use, and cultural safety
- the translation, dissemination and synthesis of evaluation findings, particularly if the OIPE was also responsible for managing the proposed Indigenous Evaluation Clearinghouse
- engaging with Aboriginal and Torres Strait Islander people, communities and organisations to facilitate their input into the Strategy’s monitoring.

It is expected that the Council would take the lead on engaging with Aboriginal and Torres Strait Islander people, communities and organisations with a focus on centring Aboriginal and Torres Strait Islander perspectives, priorities and knowledges. The OIPE would facilitate this engagement, as well as any other engagement required to enable the OIPE and Council to

draw on the evaluation knowledge and experience of organisations both within and outside government. The OIPE could provide Secretariat support for the Council.

Rewarding and applauding innovation and good practice will be an important part of the OIPE's monitoring role. This includes disseminating learnings from agencies' evaluative activities (including examples of how agencies structure their internal management of evaluation and develop capability). Innovations in evaluation, including in cases where policies or programs have not been effective, should be rewarded where it is clear that good practice evaluative standards were followed. Evaluations that provide useful insights for other monitoring and evaluation practices 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 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.

The OIPE would not be expected to undertake policy or program evaluations. As discussed earlier, evaluation should be integrated into the policy-making process, which means agencies are best-placed to undertake and commission evaluations. That said, the OIPE could identify and make recommendations to the Australian Government on, and undertake, evaluations or reviews of 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.

How should an Indigenous Evaluation Council be appointed?

In the draft background paper and Strategy, we suggested that the majority of Council members should be Aboriginal and Torres Strait Islander people. Among those who supported a Council, a majority of participants argued that all members of the Council should be Aboriginal and Torres Strait Islander people (box 7.11).

We also sought views on whether the host agency of the OIPE, or the Australian Government more broadly, should have membership on the Council. Responses to this question were mixed, with some participants expressing concerns that such members could compromise the Council's independence and impartiality. Others suggested this risk could be mitigated by limiting the number of Council members that could be appointed from government

agencies, or by including a representative from the OIPE's host agency as a non-voting member of the Council.

Box 7.11 Participants' views on membership for an Indigenous Evaluation Council

Most submissions that addressed questions about membership of an Indigenous Evaluation Council were of the view that all members (rather than a majority) should be of Aboriginal and Torres Strait Islander descent.

... the Council would be better aligned to the recognised principle and right of self-determination if all members were Aboriginal and/or Torres Strait Islander. ... It would represent a genuine commitment of the Australian government to ensuring self-determination and it would go a long way to ensuring that the overarching principle of this Strategy is delivered upon. (National Voice for Our Children (SNAICC), sub. DR163, pp. 7–8)

Consideration should also be given to a whole of Council (rather than just majority) Aboriginal and Torres Strait Islander membership to further support self-empowerment and self-determination. (Wuchopperen Health Service Limited, sub. DR117, p. 1)

If the Australian Government is genuinely committed to the overarching principle of the draft Strategy – centring Aboriginal and Torres Strait Islander people, perspectives priorities and knowledges — then all Indigenous Evaluation Council members should be First Nations people and independent of government. (Community First Development, sub. DR134, p. 10)

Aiming for all members of the Council to be Aboriginal and Torres Strait Islander people, with non-Indigenous Australian Government staff being additional to Council membership, would show the highest levels of expectation around Indigenous participation and leadership. Strong Indigenous leadership has been shown to be a key success factor for Indigenous programs ... (CSIRO, sub. DR131, pp. 6–7)

Consideration should be given to establishing a *wholly* First Nations Indigenous Evaluation Council. As an alternative, there must be processes in place that ensure majority First Nations voices are *listened to and heard* in the Council. (National Centre of Indigenous Excellence, sub. DR158, p. 12)

Views were mixed about whether the host agency should have a member on the Council.

From our perspective, there is little benefit in the host agency ... and/or Australian Government, being members of the Council. Doing so will dilute the independence and voice of the Council. (Community First Development, sub. DR134, p. 10)

The Tasmanian Government's view is that the host agency ... should also participate on the Council, as these members could provide insight into the history of previous evaluations, and what might be priorities for improvement in future evaluation work. (Tasmanian Government, sub. DR166, p. 6)

The NIAA considers the Indigenous Evaluation Council should be independent in order to provide impartial advice to the Government. One member of the [Council] could be appointed from the host agency. (National Indigenous Australians Agency, sub. DR161, p. 7)

It seems most prudent that the Office of Indigenous Policy Evaluation is co-opted as a non-voting member and that the Australian government has observer (or some similar status). Through such a structure the independence of the Council can be maintained ... (Royal Australian and New Zealand College of Psychiatrists, sub. DR138, p. 5)

The Independent Members of the National Indigenous Australians Agency's Indigenous Evaluation Committee suggested that a link to the Australian Government could otherwise be facilitated through the Council engaging regularly with the Secretaries Board, Cabinet, and individual agencies' evaluation leaders (sub. DR115, p. 3).

On how Council members might be appointed, there was broad support for Council members to represent Aboriginal and Torres Strait Islander service providers.⁵⁰ Some specified that this should be driven by the community controlled sector, while others argued that candidates from outside that sector should also be considered (Yulang Indigenous Evaluation, sub. DR150, p. 21; Alliance of First Nations Independent Education and Training Providers, sub. DR141, p. 7). The Independent Members of the NIAA's Indigenous Evaluation Council (sub. DR115, p. 3), while not specifying a particular composition of Council members, suggested that appointments should be made through a collaborative process involving the Coalition of Peaks, the Productivity Commission, and the NIAA.

There was also strong support for some (or all) members to have a research/evaluation background.⁵¹ The Tasmanian Government suggested that some members could also be appointed from relevant professional associations (sub. DR166, p. 6). Others said Council membership should reflect a diverse range of Aboriginal and Torres Strait Islander communities (including urban, rural/remote, and professional sectors), organisations, and life experiences.

On the composition of Council members, the Commission agrees that members should come from a range of backgrounds (including people with experience providing services to Aboriginal and Torres Strait people, and expertise in evaluating policies and programs), and should all be Aboriginal and Torres Strait Islander people. We propose that the Council could have up to five part-time members, including two members with expertise in evaluation, two members from the community controlled sector, and one member from the OIPE's host agency.

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 (or more accurately, as set out in chapter 6, those assessed as contributing to whole-of-Australian Government evaluation priorities and/or having a *significant* impact on the lives of 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

⁵⁰ Participants supporting Council members to represent Aboriginal and Torres Strait Islander service providers included: Community First Development (sub. DR134, p. 10); Danila Dilba Health Service (sub. DR165, p. 5); Tasmanian Government (sub DR166, p. 6).

⁵¹ Participants supporting Council members having a research and/or evaluation background included: Royal Australian and New Zealand College of Psychiatrists (sub. DR138, p. 5); National Centre of Indigenous Excellence (sub. DR158, p. 12); National Indigenous Australians Agency (sub. DR161, p. 7).

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 role, including a small team to ‘develop guidance to ensure systematic evaluation of programs and policies’, and working with the Department of the Prime Minister and Cabinet to ‘embed evaluation planning in new policy proposals’. Along with an evaluation stream within the professions model (chapter 8), 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 this approach, the Government’s response suggests a preference for incremental steps, largely building on existing efforts and practices across agencies (such as the *Enhanced Commonwealth Performance Framework*). However, there is a question about whether these changes will be sufficient to incentivise long-term changes in behaviour, or to build lasting capacity for evaluation across the APS.

... but key components are missing

The Australian Government’s 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)
- 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
- an evaluation priority-setting process at the whole-of-government level.

A more centralised approach?

Gray and Bray, in an 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. (2019, p. 25)

And Nicholas Gruen (Chair of the Open Knowledge Foundation Australia) 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 7.12 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).

Box 7.12 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) 2019).

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 7.13). 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 7.14). If such a Centre was established, the OIPE could move to the Centre as a standalone branch (figure 7.2), with the Indigenous Evaluation Council continuing its role in partnership with the OIPE. 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 results could best be disseminated to the Aboriginal and Torres Strait Islander people, communities and organisations affected by the policy or program in question.

Box 7.13 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 logics for 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 7.14 What functions could a Centre for Evaluation Excellence have?

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 an APS evaluation community of practice 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, 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 publicly on the state of evaluation across the APS.

Few participants commented on the proposal for a Centre for Evaluation Excellence, and views were mixed. The NIAA argued that the Centre, together with the OIPE, was necessary for the success of the Strategy:

The success of the Indigenous Evaluation Strategy will be dependent on the establishment of the Centre for Evaluation Excellence *and* the Office of Indigenous Policy Evaluation. Change is needed across the Australian Public Service to both evaluation capability and evaluation use to achieve the intended outcomes of the Indigenous Evaluation Strategy. Particularly important to the success of both bodies will be the appointment of an ‘eminent’ person to lead each one. (sub. DR161, p. 6)

Members of the Close the Gap Campaign Steering Committee suggested that the overarching Centre could be established first, rather than being a longer-term plan:

Less layers of administration and cost are preferred and the proposal in the background paper to establish a ... Centre for Evaluation Excellence could perhaps come at the beginning rather than the end of the process of creating an ‘evaluation culture’. (sub. DR159, p. 2)

And while not expressly addressing the Centre for Evaluation Excellence proposal, Nicholas Gruen (Lateral Economics, sub. DR177, pp. 11–12) reiterated the argument that, for true independence and objectivity, a central evaluation body must have exclusive authority to commission and conduct evaluations (rather than agencies retaining this role).

The Commission maintains the view that there is a need for independent central leadership, stewardship and oversight for social and health policy evaluations across Australian Government agencies (reiterating that evaluation should be integrated into the process of designing and implementing policies and programs). 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 an Evaluation stream within the professions model — is an important first step towards establishing a Centre for Evaluation Excellence. The Commission suggests that such a Centre be established within five years.

Figure 7.2 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: <ul style="list-style-type: none"> • Guidance material and advice to agencies on good practice evaluation governance and standards • Facilitate evaluation community of practice across APS; coordinate opportunities for upskilling in evaluative capability 	<ul style="list-style-type: none"> • Guidance and advice on Indigenous Evaluation Strategy governance and standards, and good practice • Cultural capability (so evaluations centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges) • Develop strategy for building cohort of Aboriginal and Torres Strait Islander evaluators
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 Evaluation 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

8 Evaluation capability

Key points

- Evaluation capability is 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 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 implement the Indigenous Evaluation Strategy (the Strategy) effectively.
- In the context of the Strategy, some key aspects of evaluation capability include that:
 - evaluation skills are not just important for evaluators. Those commissioning evaluations, agency leadership and APS staff need to be able to understand, plan for, and use evaluations
 - agencies take steps to strengthen their cultural capability both at the individual and organisational level, including developing strong and sustainable relationships with Aboriginal and Torres Strait Islander people
 - Aboriginal and Torres Strait Islander people's capability to engage in, partner in and lead evaluation is recognised, strengthened and supported
 - evaluations are commissioned in a way that centres Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges (and when reporting how they have upheld the principles of the Strategy, agencies should show how their commissioning arrangements have done this)
 - cultural and evaluation capability building activities need to be resourced adequately, recognising that evaluation is an investment that can lead to savings in the longer term
 - Aboriginal and Torres Strait people and organisations need to be resourced to effectively engage in evaluations.
- Agencies will be responsible for ensuring they have access to the capability — including the skills, experience and understanding — to implement the Strategy effectively. However, central support should be provided in the form of guidance, training and development.
 - The Strategy and the accompanying Guide support good practice evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people.
 - The Office of Indigenous Policy Evaluation should provide support that agencies and Aboriginal and Torres Strait Islander people can draw on to build and maintain their evaluation capability.

Evaluation capability is critical for delivering high-quality evaluations.

Evaluation capability 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 know how to effectively plan, commission and resource 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 evaluation capability in the context of 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 skills required by those undertaking evaluation (section 8.2)
- the cultural capability needed to effectively implement the Strategy (section 8.3)
- commissioning high-quality evaluations in a way that reflects the principles of the Strategy (section 8.4)
- resourcing for evaluation and capability development (section 8.5)
- roles and responsibilities for capability building (section 8.6).

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 have few staff with expertise in evaluation and/or the capability to effectively commission evaluations (or to assess the quality of commissioned evaluations).

Some participants noted a lack of capability in the APS to conduct ethical, rigorous evaluations.⁵² The NIAA, for example, said that ‘significant improvements in evaluation capability and capacity are needed’ and that:

... 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 Australasian⁵³ 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. (AES 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).

⁵² See for example: CTGSC (sub. 71, p. 6); Don Weatherburn (sub. DR114, p. 3); and SVA (sub. 83, p. 16).

⁵³ The Australasian Evaluation Society is now known as the Australian Evaluation Society.

The Commission also found that it is common practice for agencies to outsource evaluations. In the sample of evaluations given to the Commission by Australian Government agencies, only about 40 per cent were undertaken in-house (chapter 3).

Pointing to gaps in knowledge and expertise, some participants argued that there was scope to improve the cultural capability of those commissioning and undertaking evaluations (box 8.1).

Box 8.1 **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 practice 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. 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. When 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), agency leadership, 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 and programs.

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 part of everyday business (box 8.2) — needs to be embedded in 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, will require fewer technical skills than an evaluator. That said, as Gray and Bray noted, evaluation skills are also important for those commissioning evaluations.

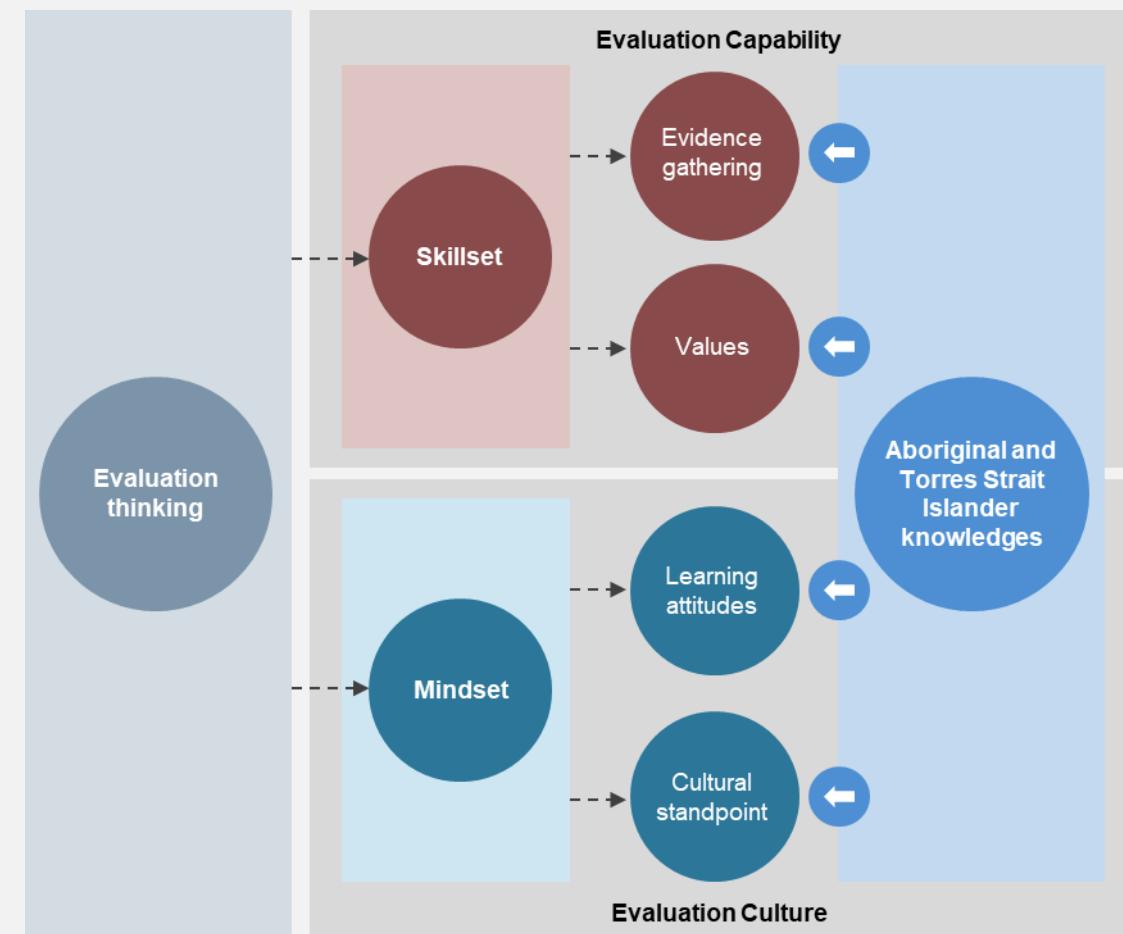
Enabling an effective evaluation function, even where this is more focused on the management of externally commissioned evaluation functions, requires staff with a suite of technical skills. These include in the area of modelling, critical to being able to undertake *ex ante* evaluation of policy options, and strong data and conceptual skills and a knowledge of evaluation approaches ... (2019, p. 25)

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.

In the context of evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, the mindset to support evaluative thinking is influenced by an agency’s evaluation culture and is fostered by 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.



Sources: adapted from Department of Education (NSW) (2020); Wehiipeihana and McKegg (2018).

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. As the NIAA said:

Our experience ... suggests that developing capability in understanding *why* ethics approvals are required is a foundational learning requirement in commissioning evaluations. (sub. DR161, p. 6)

Strengthening evaluation skills among senior managers can also help to build a culture of evaluation and ensure that evaluation is integrated into the policy cycle. As Don Weatherburn noted:

... CEOs frequently make decisions on the scale of a pilot, its duration, its location and its eligibility criteria that make it impossible to create a counterfactual or that result in a sample size too small to pick up worthwhile effects ... If the CEO and other senior managers had a basic understanding of evaluation principles and strategies, this problem could be avoided. (sub. DR114, p. 3)

A number of participants also pointed to the importance of improving the evaluation competencies of staff in the APS.⁵⁴ This is particularly important when considering evaluation at the policy design stage. As Ninti One said:

Designing a program or policy to maximise its ‘evaluability’ requires that those doing the design have a strong grasp of the principles of evaluation. (sub. DR118, p. 3)

Agency staff should be able to:

- understand what types 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 implement evaluation findings.

Undertaking evaluation requires a range of competencies

Evaluation is not a profession. There is no recognised set of qualifications for evaluators, and 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

⁵⁴ See for example: Closing the Gap Steering Committee (sub. 71, p. 6); Darling Downs and West Moreton Primary Health Network (sub. 6, p. 3); and NACCHO (sub. 95, p. 5).

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 lists of evaluator competencies that cover the skills, knowledge and experiences evaluators should have (figure 8.1). Common competencies include: the ability to undertake systematic inquiry; technical, reflective, interpersonal and project management skills; and the ability to undertake contextual analysis and engage effectively.

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 outlined some key competencies that APS employees should have 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).

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 theory Research methods and systematic inquiry	Evaluation skills	Systematic evaluative enquiry	Technical practice
Reflective		Professional foundations	Reflective practice and professional development	Reflective practice
Contextual	Culture, stakeholders and context		Contextual analysis and engagement Values and cultural competency	Situational practice
Management	Evaluative attitude and professional practice Project management	Promoting a culture of learning for evaluation Management skills	Evaluation project management and professional practice	Management practice
Interpersonal	Interpersonal skills	Interpersonal skills		Interpersonal practice

^a Australian Evaluation Society, ^b United Nations, ^c Aotearoa New Zealand Evaluation Association.

Sources: AES (2013b); ANZEA (2011); CES (2018); UNEG (2016a).

Ways to grow evaluation skills

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. There are a number of ways that this can be done.

- *Recruiting to address skill gaps* — recruitment provides a way to quickly build evaluation expertise, however, it may not lift the skills across an agency and could leave 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 8.3) — 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 and other opportunities to work directly on evaluations can give staff opportunities to develop their evaluation skills.
- *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.

8.3 Cultural capability to implement the Strategy

Cultural capability to undertake evaluations

Cultural capability is a 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). BetterEvaluation described a culturally competent evaluator as one that embraces ‘cultural and contextual dimensions important to the evaluation’, recognises their own ‘culturally based assumptions’ and takes into account ‘the differing world views of evaluation stakeholders and target communities’ (2016).

Box 8.3 Training and development for culturally capable evaluation

A number of organisations provide training and development opportunities that agencies could use to improve the evaluation and cultural capability of their staff.

- The Australian Evaluation Society provides training and development opportunities for evaluators 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 (AIATSIS) has an online cultural capability course (called the CORE program). 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 has online courses in evaluation, including graduate diploma and masters courses.
- The University of Technology Sydney offers several short online courses in evaluation.
- The Jawun program offers Australian Public Service 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.
- The Australia and New Zealand School of Government runs events and training on evaluation. It also convenes the annual First Peoples' Public Administration Conference.
- Aboriginal and Torres Strait Islander community controlled organisations can offer cultural capability training and have developed useful resources. They have specialised expertise in engaging in particular policy or geographical areas.

There are also several communities of practice — such as the Australian Public Service-wide evaluation community of practice — that provide opportunities for staff to further develop their capability.

Cargo et al. also noted that culturally safe evaluation requires evaluators to be self-aware of ‘the power they exercise in their relationships with Indigenous people, organisations and communities in all aspects of the evaluation process’ (2019, p. 2).

Cultural capability is critical for the validity of evaluation results. As BetterEvaluation put it:

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

And Weihipeihana said:

... deep cultural knowledge is needed for evaluators to work safely and respectfully with Indigenous communities ... [this knowledge helps evaluators to] respectfully navigate entry, agree on ways of working, including methods and approaches, and support Indigenous people to give expression to their views, values and experiences. (2019a, p. 374)

It is important that evaluation teams⁵⁵ are culturally capable before undertaking evaluations involving Aboriginal and Torres Strait Islander people. Working with Aboriginal and Torres Strait Islander people is key to improving evaluation teams' cultural capability. The Western Australian Government suggested that the cultural capability of an evaluator could be enhanced by:

- practitioners and commissioners working directly with Aboriginal and Torres Strait Islander people when developing and deciding on evaluation methodologies
- working with Aboriginal and Torres Strait Islander 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 Aboriginal and Torres Strait Islander 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 people in evaluations within and between government agencies (sub. 74, p. 8).

Cultural capability is also needed at an organisational level

Culturally capable evaluation teams alone will not be enough to implement the Strategy effectively. Agencies will also need to embed cultural capability within their organisations for the Strategy to be successful — without this, key elements of the Strategy (such as the Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plan (chapter 5), and priority-setting process (chapter 6)) may not be implemented well, and culturally safe evaluation may not be planned for and undertaken. One of the participants at our workshop on evaluation priority-setting commented that:

Evaluation that centres Aboriginal and Torres Strait Islander people won't go very far if the decisions made from it still occur in an Indigenous free zone.

Agencies need to go beyond seeing cultural capability as something that is 'nice to have', or only necessary for line areas that are directly involved in delivering policies and programs targeted at Aboriginal and Torres Strait Islander people. Instead, agencies need to see cultural capability as part of their 'business as usual' and as an essential prerequisite if their actions are to improve the lives of Aboriginal and Torres Strait Islander people.

Under the National Agreement on Closing the Gap (the National Agreement), Australian governments have committed to transforming government organisations, including by identifying and eliminating racism and embedding and practicing meaningful cultural safety (box 8.4). Actioning these commitments will help to embed the cultural capability necessary to implement the Strategy within agencies. And, embedding evaluative thinking will also

⁵⁵ This includes evaluators, as well as interviewers, surveyors and local researchers.

help agencies to implement these commitments by identifying what is working well and what can be improved.

Box 8.4 **Closing the Gap and the transformation of government organisations**

The National Agreement on Closing the Gap (the National Agreement), agreed to in July 2020 (chapter 1), outlines four Priority Reform Areas. The third reform area — *Transforming Government Organisations* — calls for the systemic and structural transformation of mainstream government organisations to improve accountability and better respond to the needs of Aboriginal and Torres Strait Islander people.

Under the National Agreement, governments have committed to implement the following elements within mainstream institutions and agencies:

- identify and eliminate racism
- embed and practice meaningful cultural safety
- deliver services in partnership with Aboriginal and Torres Strait Islander organisations, communities and people
- increase accountability through transparent funding allocations
- support Aboriginal and Torres Strait Islander cultures
- improve engagement with Aboriginal and Torres Strait Islander people.

To support this reform area, governments will:

- challenge unconscious biases that result in decisions based on stereotypes
- share and publish their approaches to engagement, consistent with the transformation elements
- engage with Aboriginal and Torres Strait Islander representatives before, during, and after emergencies
- include information on implementation of the transformation elements in their annual reports
- identify, develop or strengthen an independent mechanism, or mechanisms, that will support, monitor, and report on the transformation of mainstream agencies and institutions. The mechanism, or mechanisms, will:
 - support mainstream agencies and institutions to embed transformation elements, and monitor their progress
 - be recognisable for Aboriginal and Torres Strait Islander people and be culturally safe
 - engage with Aboriginal and Torres Strait Islander people to listen and to respond to concerns about mainstream institutions and agencies
 - report publicly on the transformation of mainstream agencies and institutions, including progress, barriers and solutions.

Sources: Australian Government (2020); Joint Council on Closing the Gap (2020, pp. 11–12).

Table 8.1 Organisational capability — questions to consider

Domain	Questions to consider
Governance	<ul style="list-style-type: none"> • How are Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges incorporated into agency decision making? • How are the evaluation priorities of Aboriginal and Torres Strait Islander people identified and considered? • How does the agency measure and assess the impact or effects of its work on Aboriginal and Torres Strait Islander people? • What accountability mechanisms does the agency have in place to ensure that Aboriginal and Torres Strait Islander priorities are considered? For example, are accountabilities set out in corporate plans or internal performance documents?
Relationships with Aboriginal and Torres Strait Islander people	<ul style="list-style-type: none"> • What relationships does the agency have with Aboriginal and Torres Strait Islander people? How are these relationships maintained? • How does the agency engage with Aboriginal and Torres Strait Islander people? How effective are these engagement arrangements? • What partnerships exist between the agency and Aboriginal and Torres Strait Islander people? Where can partnerships be developed and strengthened? • Does the agency have appropriate succession planning in place? How can relationships be maintained when there is employee turnover?
Structural	<ul style="list-style-type: none"> • What structures are in place within the agency to support cultural capability? Is there central support and support from agency leadership? • What mechanisms are in place to build cultural safety and address discrimination in agency processes and structures?
Workforce capability	<ul style="list-style-type: none"> • What mechanisms does the agency have in place to build workforce capability (i.e. recruitment and upskilling)? Does the agency have collaborative relationships with other organisations, such as Aboriginal and Torres Strait Islander organisations, tertiary institutions and other government agencies? • Are Aboriginal and Torres Strait Islander people represented in the agency's senior leadership?
Processes	<ul style="list-style-type: none"> • How does the agency engage Aboriginal and Torres Strait Islander people in data collection, management and use? How do data collected and used by the agency reflect the knowledges and priorities of Aboriginal and Torres Strait Islander people? Are Aboriginal and Torres Strait Islander people engaged in interpreting data? • Does the agency's policy planning consider impacts or outcomes for Aboriginal and Torres Strait Islander people? • How are Aboriginal and Torres Strait Islander people engaged in tender processes? • Does the agency have processes in place to ensure that contracted services demonstrate cultural capability? Are there processes in place to ensure the terms of contracts and assessments of value for money include requirements related to cultural capability?

Source: adapted from Te Arawhiti (Office for Māori Crown Relations) (2019c).

When seeking to improve their cultural capability for evaluation, agencies should think about how their policies and programs impact Aboriginal and Torres Strait Islander people. For example, Indigenous Business Australia has developed an impact framework that maps their programs, through a theory of change, back to a wellbeing framework. This tool provides a clear articulation of how the agency's policies and programs affect Aboriginal and Torres Strait Islander people (IBA 2020). Such a tool could be particularly useful for agencies that provide mainstream policies and programs, and who currently do not routinely assess how their policies and programs affect Aboriginal and Torres Strait Islander people.

Agencies should recognise that cultural capability exists on a continuum, and that there are always opportunities for improvement. However, given the level (and nature) of existing cultural capability within agencies varies, it is difficult to outline what the next steps are that agencies should take to improve their cultural capability. Rather, agencies need to engage in a reflective process that seeks to understand how they can better position themselves to respond to the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people. In doing so, agencies should engage with Aboriginal and Torres Strait Islander people to understand what is working well and what can be improved (while also being aware of any burdens they may be placing on Aboriginal and Torres Strait Islander people).

Partnerships and relationships with Aboriginal and Torres Strait Islander people

Strong and sustainable relationships between government and Aboriginal and Torres Strait Islander people are an important part of building cultural capability in agencies. The National Voice for our Children (SNAICC) argued that:

Delivering on [centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges] will require all levels of government to undertake a new dialogue with Aboriginal and Torres Strait Islander communities, ensuring that communities have clear pathways and opportunities to articulate their own locally led and defined aspirations. This needs to be supported in a systemic way and on a broad scale rather than in any half-hearted or tokenistic sense. (sub DR163, p. 6)

Under the National Agreement, governments have committed to working differently with Aboriginal and Torres Strait Islander people, including by ‘building and strengthening structures that empower Aboriginal and Torres Strait Islander people to share decision-making authority with governments to accelerate policy and place-based progress against Closing the Gap’ (JCOCTG 2020, p. 5). The National Agreement also outlined what shared decision making looks like (box 8.5).

One of the benefits of strong and sustainable relationships and partnerships (compared with one-off relationships developed during the course of a single evaluation) is that they also allow the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people to be incorporated at the very first stages of policy and evaluation design (that is, before any decisions are made). As Kelaher et al. said:

The ability of evaluation to orientate itself to community needs is constrained if the community has not had input into developing higher-level priorities that inform performance frameworks and the development of a program. Aboriginal and Torres Strait Islander input at any level in any part of the process is valuable; however, self-determination involves programs being oriented to the needs of Aboriginal and Torres Strait Islander people at the highest level. When this does not occur, the different players in evaluations are constrained in their ability to deliver optimal results from a community perspective. (2018b, p. 39)

And AbSec said:

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)

Box 8.5 **Shared decision making under Closing the Gap**

The National Agreement on Closing the Gap states that shared decision making is:

- by consensus, where the voices of Aboriginal and Torres Strait Islander parties hold as much weight as the governments
- transparent, where matters for decision are in terms that are easily understood by all parties and where there is enough information and time to understand the implications of the decision
- where Aboriginal and Torres Strait Islander representatives can speak without fear of reprisals or repercussions
- where a wide variety of groups of Aboriginal and Torres Strait Islander people, including women, young people, elders, and Aboriginal and Torres Strait Islander people with a disability can have their voice heard
- where self-determination is supported, and Aboriginal and Torres Strait Islander lived experience is understood and respected
- where relevant funding for programs and services align with jointly agreed community priorities, noting governments retain responsibility for funding decisions
- where partnership parties have access to the same data and information, in an easily accessible format, on which any decisions are made.

Source: Joint Council on Closing the Gap (2020, p. 6).

Longer-term relationships should also help agencies to centre Aboriginal and Torres Strait Islander people in some of the actions required under the Strategy, such as the Impact Assessment and Evaluation Plan (chapter 5) and priority setting (chapter 6).

However, while relationships may be able to drive outcomes needed to effectively implement the Strategy, they will only be able to do so if there is a mutual understanding of the needs and aspirations of Aboriginal and Torres Strait Islander people. If done poorly, in a tokenistic manner or without genuine commitment, they have the potential to damage agencies' relationships with Aboriginal and Torres Strait Islander people, and/or hinder effective decision making. SNAICC explained that:

... while partnerships have long been recognised as important for developing effective and culturally competent and safe services, many Aboriginal and Torres Strait Islander people have had negative experiences of tokenistic relationships labelled as partnerships. This has led to a level of mistrust of governments and non-Indigenous services and a belief that they may use partnerships to 'tick boxes' of cultural competence and community engagement, without a deeper commitment to sustainable relationships or local community empowerment. (2020, p. 4)

This points to the importance of ensuring that agencies demonstrate a commitment to building genuine relationships with Aboriginal and Torres Strait Islander people.

Improving cultural capability by using Aboriginal and Torres Strait Islander capability

Ernst & Young suggested that:

The deep insights and intelligent cultural nuancing First Nations' evaluators can bring would play a fundamental role in improving the quality of policy and programs, as well as signalling the value government places on this critical aspect of evaluation. (sub. DR140, p. 5)

Along a similar line, Wehipeihana argued that while many things can be learned about culture, 'some things can be known and revealed only from within the culture' and for this reason 'there is no substitute for the cultural capital (understanding, knowledge, and intuit) that comes from being of the culture' (2019a, p. 373).

Wehipeihana also noted that evaluations led by Indigenous people are more likely to have the cultural capital to:

- facilitate respectful engagement and observe 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).

However, because the current pool of Aboriginal and Torres Strait Islander evaluators is small, agencies may not be able to secure the services of an Aboriginal and/or Torres Strait Islander evaluator with the capabilities that match what the evaluation requires.

Where this is the case, consideration should be given to how Aboriginal and Torres Strait Islander people can be involved in the evaluation — that is, beyond being evaluation participants — so that Aboriginal and Torres Strait Islander voices can be heard. For example, an Aboriginal and/or Torres Strait Islander evaluator and a non-Indigenous evaluator could provide, in combination, the technical and cultural expertise required to conduct an evaluation well.

Some participants argued there was a need to build up the pool of Aboriginal and Torres Strait Islander evaluators. 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)

And Don Weatherburn said:

Evaluation design requires certain technical skills that, at present, are in very short supply amongst Indigenous university graduates. There is a pressing need to increase the number of Indigenous Australians with tertiary level training in evaluation research. (sub. DR114, p. 5)

The requirement under the Strategy for agencies to centre Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges could increase the demand for culturally capable evaluation (and opportunities for Aboriginal and Torres Strait Islander evaluators to lead evaluations). It could also provide opportunities for more Aboriginal and Torres Strait Islander people to gain skills and experience in other aspects of the evaluation process, such as deciding on evaluation design, undertaking data collection and interpreting results, which should expand the existing pool of Aboriginal and Torres Strait Islander people who can lead evaluations. Participants⁵⁶ also suggested other ways to strengthen the capability of Aboriginal and Torres Strait Islander people to lead evaluation, including by:

- 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.

These suggestions all have merit. In particular, the Commission considers that there is significant opportunity to strengthen Aboriginal and Torres Strait Islander people’s capability to lead evaluation by:

- growing a cohort of Aboriginal and Torres Strait Islander evaluators within the APS

⁵⁶ 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).

-
- governments working better with service providers and providing funding for continuous quality improvement and evaluation.

Growing a cohort of Aboriginal and Torres Strait Islander evaluators in the APS would offer a number of benefits, all of which would assist with the Strategy's implementation.

More Aboriginal and Torres Strait Islander evaluators would make it easier for agencies to have Aboriginal and Torres Strait Islander people lead evaluations. A network of Aboriginal and Torres Strait Islander evaluators could also help shift the evaluation culture in the APS to focus more 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, in 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)

A strategy should be developed to build a cohort of Aboriginal and Torres Strait Islander evaluators within the APS. As part of developing this strategy, a range of actions should be considered, including:

- actions to grow the number of Aboriginal and Torres Strait Islander evaluators working in agencies such as streamlined or centralised recruitment processes
- actions to strengthen the skills, experience and capabilities of evaluators, such as providing structured training, networking opportunities, mentorship programs and secondment opportunities.

This strategy should be developed by the Office of Indigenous Policy Evaluation (OIPE) (chapter 7).

Working with service providers for quality evaluations

Working better with those who deliver policies and programs — particularly Aboriginal and Torres Strait Islander community controlled organisations — could also help lift the quality of evaluations. Because service providers provide 'frontline' services they can provide valuable insights on what is happening on the ground, collect and provide data, act as a contact point between evaluators and service users and implement evaluation recommendations. Community controlled organisations can play a particularly valuable role in evaluations, providing both connections to the community and representing Aboriginal and Torres Strait Islander people (box 8.6).

The National Agreement recognised the importance of the community controlled sector, noting that 'an Aboriginal and/or Torres Strait Islander Community-Controlled Organisation delivers services, including land and resource management, that builds the strength and empowerment of Aboriginal and Torres Strait Islander communities and people' (JCOCTG 2020, p. 8).

Box 8.6 **Comments on the role of service providers in evaluations**

Victorian Aboriginal Child Care Agency (VACCA):

VACCA's experience has demonstrated clearly that evaluations by government funding agencies draw significantly on service providers in implementation ... A common scenario for VACCA is that a government agency funding a program will commission the evaluation, use existing program data (both administrative and monitoring), and either conduct additional data collection activities or require that VACCA collect additional data to support the evaluation (surveys, for example). (sub. DR126, p. 2)

PwC Indigenous Consulting:

Where these organisations have strong connections with their wider community, they provide access to crucial health and human services; have strong and trusted community relationships; and could be a valuable resource for data collection and evaluation. (sub. 58, p. 3)

Lowitja Institute:

Investing in the capacity of community-controlled organisations to develop and house their own evaluation processes has the potential to look at models of evaluation that:

- inform localised decision making
- encourage localised (or community led) policy and program cycles
- increase social capital
- support meaningful partnerships and co-design processes
- facilitate community driven Knowledge Translation and Data Sovereignty. (sub. DR157, p. 3)

However, the way governments commission and fund service providers can affect how providers contribute to evaluations. The Commission has previously found that:

- reporting requirements are not always relevant for informing how to improve a service
- reporting requirements can be burdensome for service providers
- government statistics and data are reported at a national or jurisdictional level rather than at the community level, limiting their usefulness for both evaluation and service providers
- funding for service providers may not recognise that evaluation and activities to improve service quality are an important part of providing services (PC 2017b, chapter 8, 2020, chapter 9).

Requiring agencies to plan early for evaluation and to engage early with evaluation participants is one way that the Strategy will improve how government works with service providers to evaluate policies and programs. In practice, this means agencies engaging with service providers to define what success looks like for them when deciding on evaluation design and data collection processes. SNAICC identified some key practices for monitoring and evaluating in partnership with community controlled organisations (box 8.7).

Box 8.7 Key practices for monitoring and evaluating in partnership with community controlled organisations

1. Indicators of success are mutually agreed between partners and, where appropriate, jointly reported on.
2. Where a community controlled organisation reports to a partner who is also a funder, the community controlled organisation participates in developing reporting processes which are relevant and not too onerous.
3. Partners provide support to community controlled organisations for evaluation capacity development and data collection processes.
4. Community controlled partners participate in the design of evaluation and review processes.
5. Evaluation processes acknowledge and incorporate Aboriginal cultural perspectives on evaluation including, for example, qualitative feedback and storytelling approaches.

Source: SNAICC (2012, pp. 51–52).

However, it is important that as part of early planning, agencies consider what resources are needed for a service provider to effectively evaluate and engage in continuous quality improvement activities. Requirements for monitoring and evaluation should be embedded into service contracts and appropriately resourced. This should strengthen the capability of service providers to undertake evaluations and improve the quality of evaluations.

8.4 Capability to commission high-quality evaluations

Current commissioning practice

As discussed in chapter 2, it is common for agencies to commission evaluations to external organisations, and because building in-house expertise to conduct evaluations will take time (and for some smaller agencies, maintaining strong in-house evaluation expertise may not be viable), agencies are likely to continue to rely predominately on commissioning evaluations to undertake evaluations under the Strategy.

However, 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)

And Kimberley Aboriginal Medical Services said it had seen:

... a spate of evaluations for national Australian Government health projects in which consultants are the evaluators. The competitive market tendering of evaluations pays lip service to Aboriginal involvement in the process but has poor recourse to local capacity building, local contextualization of data and does not contribute to work opportunities for local Aboriginal people. (sub. DR164, p. 2)

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. (Lowitja Institute, sub. 50, p. 6)

Kelaher et al. (2018b), when they reviewed a 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, 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 to this question, but there are a number of factors that should be considered when deciding who is best 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 participants 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 evaluation is about learning and improving the way a policy or program is implemented. 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). Lateral Economics called evaluation consultants ‘independence-for-hire’ (sub. DR177, p. 12).

For these reasons, evaluation transparency and oversight are critical to provide 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 policy and program design and delivery.

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

- 1 Establish an evaluation advisory group that includes stakeholder representatives to guide and inform the evaluation process.
- 2 Develop a program logic model (ideally developed in program planning) to explain the causal pathways linking program activities, outputs, intermediate impacts and longer term outcomes.
- 3 Develop an evaluation plan that includes:
 - overview of the program
 - purpose of the evaluation
 - audience for the evaluation
 - evaluation questions
 - evaluation design and data sources
 - potential risks
 - resources and roles, including budget and timeline
 - governance
 - reporting.
- 4 Procure an independent evaluator, by:
 - preparing a request for quote or request for tender
 - issuing an invitation to quote or tender
 - engaging an evaluator and agreeing to a contract.
- 5 Project manage the development and implementation of the evaluation workplan and the achievement of the contract milestones.
- 6 Disseminate the evaluation findings to support the incorporation of results into program decision making.

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 8.8).

Box 8.8 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 (maintained by the Department of Social Services) and the Collaborative Evaluation, Research and Planning Panel (maintained by the NIAA). 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, there are a number of things 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 must⁵⁷ 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

⁵⁷ Under Commonwealth procurement rules (Department of Finance 2019, pp. 11–12).

also reflect any ethical requirements, such as the treatment of intellectual property, ethics clearance and engagement requirements.

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.

Once 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 evaluators or 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.

Commissioning evaluations and centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges

Both commissioners and evaluators are responsible for ensuring that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred in commissioned evaluations. As Kelaher et al. noted, 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, resourcing 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. Table 8.2 sets out questions for commissioners to consider.

Table 8.2 Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges — questions for commissioners

Commissioning steps	Questions to consider
Planning for evaluation	<ul style="list-style-type: none"> • How are Aboriginal and Torres Strait Islander people engaged in evaluation planning? <ul style="list-style-type: none"> • For high priority evaluations, agencies should ensure evaluation governance includes Aboriginal and Torres Strait Islander representation. • How do evaluation questions and outcomes reflect what is important to Aboriginal and Torres Strait Islander people? • What timelines and resources are needed for Aboriginal and Torres Strait Islander people to effectively engage in the evaluation?
Commissioning the evaluation	<ul style="list-style-type: none"> • 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? <ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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? <ul style="list-style-type: none"> • For high priority evaluations, agencies may want to require results to be translated either by an Aboriginal and/or Torres Strait Islander evaluator or by a non-Indigenous evaluator in partnership with Aboriginal and Torres Strait Islander people, communities, or organisations. • How are management responses communicated back to the community and organisations?

Agencies also need to consider how their commissioning processes centre 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 8.9. 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 about how to address the questions raised in table 8.2. Box 8.10 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, there are many ways evaluations of mainstream policies and programs 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 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 8.9 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 — that the proposal be both technically and culturally appropriate; that key personnel undertaking the project can demonstrate research and evaluation experience with Victorian communities; and that the bidder had recent work experience and existing relationships with the communities engaged in the evaluation.
- The evaluation team had three researchers, of which two 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 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 8.10 A hypothetical commissioned evaluation of a business support program

A hypothetical business coaching program aims to help small to medium business enterprises expand. It is not targeted at Aboriginal and Torres Strait Islander businesses. Program managers want to know how to improve the program for users, including Aboriginal and Torres Strait Islander people.

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.

The 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.

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.

8.5 Resourcing evaluation

Evaluation needs to be sufficiently funded if it is to be done well. As discussed in chapter 5, the benefits of evaluations are more likely to be realised when results are rigorous, users and providers are engaged throughout the process and evaluation findings are presented in relevant and accessible ways — all of which require adequate funding.

There will also need to be investments in evaluation capability to support a culture of learning and evaluation in the APS — something that project participants identified:

The Strategy, if appropriately implemented, will provide useful guidance to Commonwealth departments in planning and conducting evaluations. However, the proposed actions and principles outlined in the Strategy will require a major shift in the culture and processes of Commonwealth departments, and significant time and resources will be needed to ensure that agencies are equipped with the required skills and expertise. (AMA, sub. DR116, p. 1)

The Federal government must do more than just have a strategy in place. Departments and service providers must be resourced adequately to undertake evaluation processes properly ... One of the likely benefits of good evaluation processes is the ability to identify value. However, an underfunded Public Service will not be able to deliver in line with the quality of the [Indigenous Evaluation Strategy] standards. (ANTaR, sub. DR124, p. 5)

That the Strategy is supported by increased investment in targeted cultural competency training and capacity strengthening of key government agencies to build effective and trusted relationships with Aboriginal and Torres Strait Islander stakeholders, sectors and communities as a critical starting point. It cannot be assumed that success and improvements will occur organically, within existing resource constraints or by simply applying greater effort or focus on it as a mere objective. (SNAICC, sub. DR163, p. 5)

As outlined in chapter 2, the Commission heard from agencies that funding is often cut from New Policy Proposals and, as such, agencies are often left to fund evaluation activities through their budget appropriations. This was cited as a reason why evaluation is not undertaken. While it is difficult to assess how common this is, what is apparent is that some agencies have managed to develop evaluation systems and undertake a significant amount of evaluation under current funding arrangements.

This suggests that evaluation resourcing is a question of priorities as well as one of funding availability. So while providing part of the funding for evaluation through New Policy Proposals could be beneficial, agencies should not use a lack of earmarked funding as a reason for not undertaking evaluation.

Agencies need to recognise that evaluation is an essential part of the policy cycle and resourcing needs to reflect this. Without evaluation, improvements to policies and programs may not be realised. Evaluation should be seen as an investment, rather than as an impost or cost for no real benefit. As the NHLF said:

... agencies will require investment and support to develop their technical and cultural skills and capability, which is in addition to having Aboriginal and Torres Strait Islander people involved throughout the process. ... this type of investment will bring better results and outcomes in the future, which are likely to create savings in the long term. (sub. DR146, p. 4)

Resourcing engagement with Aboriginal and Torres Strait Islander people

Another common theme the Commission heard was that Aboriginal and Torres Strait Islander people are often not resourced to participate in evaluation (box 8.11) but rather, there is an expectation that Aboriginal and Torres Strait Islander people and organisations will give up their time to participate in evaluations because it is in their interest to do so. Some also pointed out that Aboriginal and Torres Strait Islander people and organisations were expected to lend cultural expertise to improve the credibility of an evaluation without adequate compensation for their contributions. As outlined above, there are also challenges for service providers and community controlled organisations to participate in evaluations when they have not been funded to do so.

Box 8.11 Participants' views on resourcing Aboriginal and Torres Strait Islander engagement

Liz Curran:

Often the agencies being evaluated have little resources to do the service provision and are expected to drop everything by those conducting or commissioning the evaluations so as to feed into and support the external evaluation with the threats to funding or sustainability underpinning the evaluation. (sub. 2, p. 3)

Victorian Aboriginal Community Controlled Health Organisation:

If there is an expectation that Aboriginal people and Communities are to provide their knowledge and expertise in evaluating policies and programs, they must be appropriately reimbursed financially to accommodate for lost time for individuals or organisations providing service delivery. (sub. 44, p. 19)

First Nations Media Australia:

... First Nations Media Australia is supportive of an evaluation approach that includes Aboriginal and Torres Strait Islander people, through community-controlled peak bodies informed by our members, in the design of evaluation methods and the collection and analysis of data. That said, this work needs to be appropriately funded. It should not be expected that First Nations organisations can absorb participation in evaluation processes within their current operations, or that First Nations organisations should carry the additional burden of providing cultural competency support to non-Indigenous consultants and researchers without reasonable compensation for that expertise. (sub. 30, p. 13)

National Justice Partnership:

If government genuinely wants to include Aboriginal and Torres Strait Islander organisations and peoples in the evaluation and review stages of policy, program and service planning, it is essential that clear and consistent funding is allocated for this purpose, not just 'within agencies' but within funded organisations. (sub. 51, p. 9)

ANTaR:

As a signatory to the [United Nations Declaration on the Rights of Indigenous Peoples], Australia is obligated to support Aboriginal and Torres Strait Islander peoples' rights of self-determination ... and their capacity and capability to properly engage with the [Indigenous Evaluation Strategy] process. (sub. DR124, p. 6)

Victorian Aboriginal Child Care Agency:

... government agencies need to adequately fund service providers and community organisations and representatives for their engagement in monitoring and evaluation activities. (sub. DR126, p. 5)

The Strategy's overarching principle of centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges positions Aboriginal and Torres Strait Islander people to be much more than just participants in an evaluation. It asks that governments and Aboriginal and Torres Strait Islander people work together to produce high-quality evaluations that are relevant to improving the lives of Aboriginal and Torres Strait Islander people. However, Aboriginal and Torres Strait Islander people and organisations need the resources to be partners with government.

The National Agreement spoke to this, noting that 'adequate funding is needed to support Aboriginal and Torres Strait Islander parties to be partners with governments', including to:

- engage independent policy advice
- meet independently of governments to determine their own policy positions

-
- support strengthened governance between and across Aboriginal and Torres Strait Islander organisations and parties
 - engage with and seek advice from Aboriginal and Torres Strait Islander people from all relevant groups within affected communities, including but not limited to Elders, Traditional Owners and Native Title Holders (JCOCTG 2020, p. 7).

Given this, agencies will need to consider the resources needed for Aboriginal and Torres Strait Islander people and organisations to effectively engage in evaluations and evaluation processes.

8.6 Who should be responsible for building evaluation capability?

Agencies should be responsible for building their own capability ...

The Strategy places a range of requirements on, and provides a range of guidance to, agencies to improve how they evaluate policies and programs affecting Aboriginal and Torres Strait Islander people. Agencies will need to ensure that they have access to the skills, resources and capabilities they need to meet these requirements and to apply this guidance. Placing the onus for capability building on agencies means that they will have the freedom to invest in the activities that provide them the greatest benefit based on their needs and their current capabilities.

Agencies will be best placed to determine what capability they need and the best way to develop it. The types of capabilities that agencies will require will vary, 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).

... with some central support

While responsibility for developing and maintaining evaluation capability should lie with agencies, a central function should 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* (the Guide) 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 8.12), the Australian Government also agreed to establish an evaluation profession.

Box 8.12 The Independent Review of the APS professions model

The Independent Review of the Australian Public Service 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. Digital and data professions were also launched in 2020.

Sources: APSC (2019, 2020a, 2020b, 2020c, 2020d); DPMC (2019d, pp. 20, 22, 2019e, pp. 200, 223).

In the draft Strategy, the Commission said that the Head of Evaluation Profession should have responsibility for building evaluation capability. However, it is not clear when the evaluation profession will be implemented. Due to the COVID-19 pandemic, the Australian Government paused its implementation planning of the reform agenda developed in response to the Independent Review of the APS (Gaetjens and Woolcott 2020b), and the Secretaries Board has since indicated that it would prioritise initiatives that ‘build on and embed the best of the APS during the crisis’ (Gaetjens and Woolcott 2020a). It is not clear whether the evaluation profession will be prioritised.

Given this uncertainty, responsibility for supporting agencies to build evaluation capability should be placed with the OIPE, at least in the short term. This means that the OIPE would have a wider remit than the Commission proposed in the draft Strategy. However, the

Commission considers that the OIPE would have the necessary features to support a capability-building function. There would also be benefits in housing the capability function alongside the other functions of the OIPE to help share learnings and good practice.

In undertaking this capability function, the OIPE should:

- 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.

As mentioned above, the OIPE should also develop the strategy to grow the cohort of Aboriginal and Torres Strait Islander evaluators.

The OIPE should leverage the expertise of other agencies and Aboriginal and Torres Strait Islander organisations to undertake this capability-building function. This will help to prevent duplication of efforts, and to share lessons on how to build capability most effectively. The OIPE should also have training and development opportunities that are open to Aboriginal and Torres Strait Islander people, including those who work outside the APS. This would facilitate two-way learning between government and Aboriginal and Torres Strait Islander people and strengthen the capability of Aboriginal and Torres Strait Islander people to engage in, partner in, and lead evaluations.

There is likely to be a degree of overlap between the capability-building function of the OIPE and the functions of the Head of Evaluation Profession (once the latter is established). The proposed review of the Strategy (chapter 10) — to occur five years after the Strategy is first endorsed — would provide an opportunity to examine any overlap, and recommend ways for this to be addressed.

Monitoring the Strategy will provide an opportunity to assess agencies' efforts to build their capability

Even with clearer expectations and access to support, there remains a risk that without the right incentives, agencies could underinvest in building their evaluation capability. An important part of monitoring the Strategy (chapter 10) will be examining agencies' efforts to develop and maintain evaluation capability to meet their responsibilities under the Strategy. By assessing the quality of evaluations, the monitoring arrangements for the Strategy will be able to identify evidence of gaps in the evaluation skills and/or the cultural capability of agencies. The monitoring process will also provide a check on whether agencies are commissioning evaluations in a way that is consistent with the Strategy and are adequately resourcing evaluations.

The monitoring process should help to lift the evaluative capability of the APS, and others. It should involve assessment of the guidance provided under the Strategy, including whether further guidance is needed. And while the monitoring process should focus on the conduct of agencies, the good practice and lessons identified could also be applicable to others, such as Aboriginal and Torres Strait Islander evaluators, communities and organisations as well as other governments and non-government organisations.

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 (through surveys, interviews or focus groups). A review of recent Australian Government agency evaluations, undertaken by the Commission, showed that most collected new data rather than using monitoring or other existing data.
- There were a number of concerns raised about data for evaluation, including:
 - data burdens on participants and service providers
 - data content, including a lack of data that reflect the priorities of Aboriginal and Torres Strait Islander people, and data that focus on outputs rather than outcomes
 - poor data quality, such as inconsistent or missing data.
- 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.
- Engaging with Aboriginal and Torres Strait Islander people on decisions about data collection, management and use will: ensure that collected data reflect what is important to Aboriginal and Torres Strait Islander people; improve trust in data collection and governance arrangements; ensure that data are collected, managed, used and stored in an ethical and culturally appropriate way; and support Aboriginal and Torres Strait Islander data sovereignty.
- Evaluation data will also be improved by:
 - addressing data gaps and greater consistency of data collection across agencies
 - 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)

Early planning for what data should be used 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. The Magenta Book noted that:

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)

Engaging with Aboriginal and Torres Strait Islander people on decisions about data collection, management and use at the early planning stage is important for ensuring data reflect what is valued by Aboriginal and Torres Strait Islander people and communities. As Maggie Walter said:

Good evaluation requires good data. If the right questions and evaluation logics are put in place, and Indigenous input is there from the start, then good data processes will ensue. (sub. 112, p. 4)

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 and to provide context to an evaluation.

Box 9.1 Sources of data

There are many different sources of data that can be used for an evaluation.

- *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. They generally cover all aspects of a policy or program including inputs, processes, outputs and outcomes. A particularly useful feature of monitoring data is that they provide information 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* — large-scale surveys can 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 be collected as new data, although some monitoring data can be qualitative. A variety of methods may be employed to collect qualitative data including interviews, focus groups, case studies and observations (chapter 4 looks at the strengths and weaknesses of the different methods).

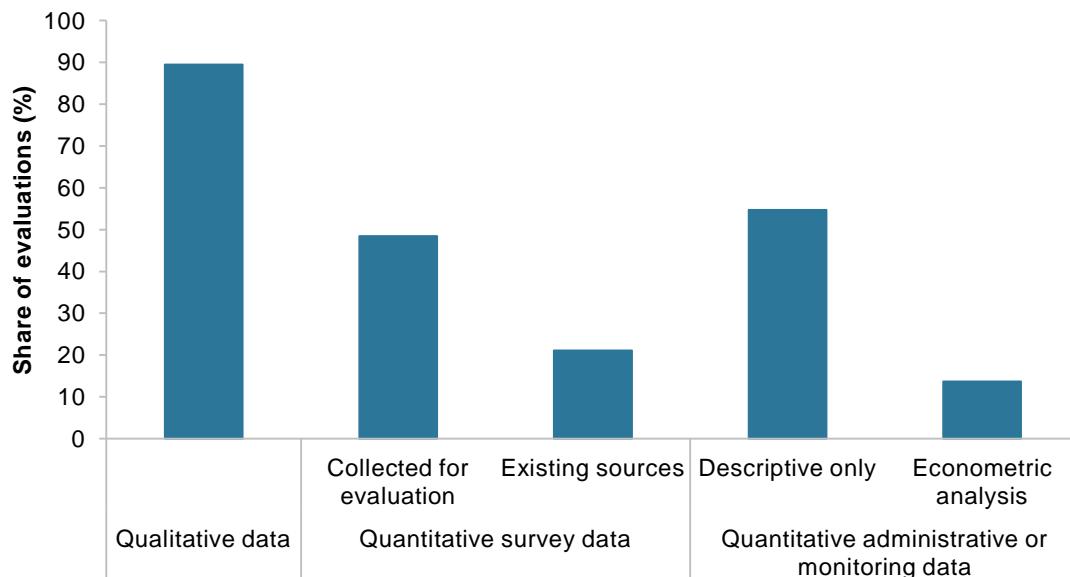
Source: adapted from HM Treasury (2020c).

Evaluations rely heavily on new data

As discussed in chapter 3, most of the evaluations conducted or commissioned by Australian Government agencies and provided to the Commission via our information request (appendix B) used both qualitative and quantitative data. Most evaluations collected new data rather than using existing data from surveys, administrative or monitoring data (figure 9.1).

Figure 9.1 Data sources used in recent Australian Government evaluations^a

Per cent of evaluations that report using each data type^b



^a Based on a review of 96 evaluation reports submitted to the Commission as part of the information request to Australian Government agencies. The sample includes evaluations of both Indigenous-specific policies and programs and mainstream policies and programs that examined impacts on Aboriginal and Torres Strait Islander people. ^b Most evaluations report using multiple data types.

Source: Productivity Commission analysis. See Appendix B for further details.

Quantitative data used in the evaluations were from a range of sources. There was a greater reliance on existing data sources for quantitative data 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.

- Almost 50 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.
- 21 per cent of evaluations used data from existing quantitative surveys. These included regular surveys of providers and jobseekers undertaken by the Department of Employment (now the Department of Education, Skills and 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 14 per cent of evaluations used existing administrative or monitoring data for econometric analysis or inferential statistics.

9.2 Issues raised about data for evaluation

Data collection and analysis can be costly and burdensome

There are costs — to government agencies, program providers and participants — to collecting 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 Australia:

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)

NSW Aboriginal Land Council:

NSWALC agrees with the Coalition of Peaks that without stronger commitments on how evaluations will be used by Governments, there is a risk the Indigenous Evaluation Strategy will add to the burden of reporting many Aboriginal Community Controlled organisations currently face, without delivering any benefits. (sub. DR122, p. 3)

A recent review of Indigenous-specific primary health care datasets found that some organisations are required to complete 50 reports a year. The review 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 data providers when data are not used by government, are not useful to the data providers themselves, or do not contribute to improved outcomes. The 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)

Ernst & Young saw the Indigenous Evaluation Strategy (the Strategy) as an opportunity to improve the usefulness of the data collected for evaluation:

Historically, in the context of evaluation, communities have been exploited with little positive benefit arising from use of their data. The [Indigenous Evaluation Strategy] now has the opportunity to re-define how communities become empowered to be custodians of their data, on country, where knowledge and data is accessible and applied in ways that best support the aspirations of those communities. (sub. DR140, p. 6)

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).

Where this is the case, indicator improvements 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. As an example, 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 [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 [health justice project] evaluation context. (sub. 18, p. 4)

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)

Victorian Aboriginal Child Care Agency:

Program logics developed by funding agencies often reflect the outcomes being sought by funding agencies — which might relate to broad policy reforms rather than specific intended outcomes for service users. They might also not include cultural outcomes that matter to service users and service providers in achieving change. (sub. DR126, p. 4)

National Voice for our Children (SNAICC):

[Co-designing methodologies] includes the opportunity to define the outcomes and measures that are important to Aboriginal and Torres Strait Islander people, rather than only responding to government defined goals and targets. (sub. DR163, p. 7)

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, for example, 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 tried to improve the collection of data relevant to Aboriginal and Torres Strait Islander people's lives. Examples include the *Mayi*

Kuwayu study and the *Community Wellbeing from the Ground Up: A Yawuru example* research project (Jones et al. 2018; Yap and Yu 2016a). The data collected in these projects cover topics such as cultural connectedness and identity, family relationships, experiences of racism, cultural safety and voice.

The National Agreement on Closing the Gap (the National Agreement) also outlined areas for data development to better capture drivers, contextual information and outcomes that are important to Aboriginal and Torres Strait Islander people (box 9.4). Data developments include measures of service access, experiences of racism, and availability of culturally safe services, among others.

Monitoring data are often focused on outputs or compliance rather than outcomes

While data on inputs and outputs are important, outcome data are needed to assess impact. However, many evaluations use data that focus heavily 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 is not planned during policy and program design and is 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).

Box 9.4 Closing the Gap outcomes and data development

The National Agreement on Closing the Gap (the National Agreement) has two sets of targets to assess progress against the objectives and outcomes of the agreement. These are:

- Priority Reform targets that focus on measuring changes that governments are making in the way that they work with Aboriginal and Torres Strait Islander people
- socio-economic targets that focus on measuring the outcomes experienced by Aboriginal and Torres Strait Islander people.

The National Agreement acknowledged that the full implementation of the Priority Reforms will support the accelerated achievement of the socio-economic targets.

There are 16 socio-economic targets, and each contains:

- outcomes — the socio-economic outcomes of the National Agreement
- targets — measures that hold governments publicly accountable. Each includes a measurable goal to be met over the next decade
- indicators — supporting measures to provide a greater understanding of how governments are tracking. They identify drivers and contextual information. Drivers measure the factors that significantly impact the progress made toward a target. And contextual information provides insight into the experiences of Aboriginal and Torres Strait Islander people under each outcome
- disaggregation — how the reporting of each target will be broken down
- data development — areas that are important for understanding outcomes, but cannot be measured currently.

The National Agreement also committed to a data development plan being established for endorsement by the Joint Council on Closing the Gap within two years of the National Agreement commencing. The data development plan will:

- be developed in partnership and be agreed to by all parties
- prioritise data development actions over the life of the National Agreement
- outline clear timeframes for actions to be delivered and who will be responsible for each action
- be reviewed by the Joint Council at the same time as it reviews the Productivity Commission and Aboriginal and Torres Strait Islander-led reviews, at which time the Joint Council may consider changes to the plan.

Source: Joint Council on Closing the Gap (2020).

The NSW Aboriginal Housing Office suggested that part of the problem stems from commissioners of evaluations and service providers being more familiar with collecting data on program outputs and key performance indicators (KPIs):

... there is less understanding of the outcomes data that may be required and the importance of consistently collecting such data to avoid data gaps and measure the impacts of programs by establishing baselines. (sub. DR130, p. 2)

Considering outcomes as well as outputs in program KPIs can help ensure that monitoring data are more useful for evaluation and can reduce reporting burdens. The NSW Aboriginal Land Council noted that:

... there is scope to change the way in which KPIs for government funding agreements are developed to ensure that the information organisations are asked to report on reflect meaningful outcomes ... The KPIs should be based on evaluation program logic and involve asking meaningful questions. Ensuring that organisations are reporting on relevant KPIs will help to incorporate evaluation methodologies into the funding of programs and services without overburdening organisations with unnecessary, additional reporting requirements. (sub. DR122, p. 4)

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
- help identify duplication, as well as returns on investment.

They are also important for assessing what policies and programs should be evaluated and whether there are opportunities for cross-cutting evaluation.

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 agencies will have information about the policies and programs they deliver, they may not have data on policies and programs delivered by other Australian Government agencies, state, territory and local governments or non-government organisations. A lack of data and transparency about policy and program inputs can also hinder community-led or place-based evaluation efforts.

The high-level estimates of expenditure on Aboriginal and Torres Strait Islander people that the Commission produces for the Steering Committee for the Review of Government Service Provision have limited usefulness for evaluation because:

- 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 categories of expenditure are too broad to usefully inform evaluations, and data are not available at the policy or program level.

The Commission's study on *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 establishing the causal impact of policies and programs (chapter 4). Decisions about what baseline and comparison data to collect need to be made in the early planning stage of an evaluation and with Aboriginal and Torres Strait Islander people. It is important that baseline and comparison indicators reflect the outcomes that will be measured after implementation.

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)

It is common for evaluations to identify a lack of baseline or comparison group as a significant limitation (AHA 2018; KPMG 2019; Mitchell, Bandara and Smith 2018) and often it is a lack of pre-program data that is the cause of the problem. For example, trial data from the Department of Social Services' Individual Placement and Support trial only covered the final two years of the three-year trial (KPMG 2019). And the Cashless Debit Card evaluation commissioned by the Department of Social Services in 2016 was commissioned only a few weeks before the program was implemented, which meant that there was not enough time to collect adequate baseline data. Commenting on the evaluation, 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. 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)

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 Department of Jobs and Small Business' Relocation Assistance Program, a comparison group could not be identified as there were differences between those who decided to move and those who did not, and these could have influenced outcomes (DJSB 2018).

Data collected 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.

Some of the problems with poor-quality monitoring data stem from evaluation data needs not being considered when policies and programs are planned and implemented. Don Weatherburn suggested that ‘because senior officers often have a poor understanding of outcome evaluation strategies, they often fail to establish the data systems needed to make rigorous evaluation possible’ (sub. DR114, p. 4).

The quality of monitoring data can be improved by ensuring that program administrators and service providers understand how monitoring data will be used to improve policies and programs, as well as allowing sufficient resources for data collection and management. As the NSW Aboriginal Housing Office explained:

Given the historical research experiences of Aboriginal and Torres Strait Islander people, it is even more important to have transparency around such data requirements when engaging Aboriginal-led or controlled organisations and have evaluation experts on hand to explain how the findings from such data can be used to benefit Aboriginal people. If such data requirements are not outlined from the outset and embedded in contractual obligations, given the historical research experiences of Aboriginal and Torres Strait Islander people, it is understandable that they may hesitate to collect some data. (sub. DR130, pp. 2–3)

Box 9.5 **Limitations of monitoring and administrative data identified in recent Australian Government evaluations**

An Evaluation of the Family Advocacy and Support Services

Some key variables were not available in the administrative data. Differences in data collection methods also made comparisons across jurisdictions difficult.

Tackling Indigenous Smoking program: Final Evaluation Report

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

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

There were inconsistencies in the way service providers collected and reported on data in quarterly reports.

Settlement Grants Program

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

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

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

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

Some outcome measures changed when a new employment services model was introduced, reducing data comparability over time.

Sources: AHA (2018); DESSFB (2020); DJSB (2019); Inside Policy (2018); KPMG (2019); Mavromaras et al. (2018); Mitchell, Bandara and Smith (2018); Smyth et al. (2017); Social Compass (2019).

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). The Tasmanian Government, commenting on the challenges faced in jurisdictions with small Aboriginal and Torres Strait Islander populations, said:

Statistical analyses of data-sets specific to the Aboriginal and Torres Strait Islander population in Tasmania have limited power and reliability, and are often not useful representations of progress and outcomes. (sub. DR166, p. 2)

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. The Smith Family argued that:

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)

The National Agreement highlights the importance of being able to disaggregate data to reflect the diversity of Aboriginal and Torres Strait Islander people. Noting that disaggregation ‘allows us to understand where progress is being made and where greater effort is needed’ (JCOCTG 2020, p. 16). Many of the socio-economic outcomes in the National Agreement will be disaggregated by remoteness, gender and age, among other characteristics.

Small 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 some samples (AHA 2018; Social Compass 2019).

Sample sizes are also important when considering data that are not solely collected for evaluation. Some existing data sources have small Aboriginal and Torres Strait Islander sample sizes or experience challenges with attrition. 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 response rates and increase sample sizes. The success of the Longitudinal Study of Indigenous Children (LSIC) is notable in this context. 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 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.

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. (2013, p. 6)

Many factors influence a person's choice to self-identify. The ABS and Aboriginal Affairs (NSW) 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/or 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/or 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

^a This is a document confirming an individual's Aboriginal and/or Torres Strait Islander heritage.

Sources: Aboriginal Affairs (NSW) (2015); ABS (2013, 2018).

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 National Indigenous Reform 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). 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.

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.

9.3 Improving data for evaluation

There are a number of ways data collection, quality and governance could be improved. 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
- improving the consistency of data collection across agencies
- making better use of existing data in evaluation.

Planning early for evaluation and data collection

The benefits of early planning

Many data problems stem from poor data planning 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)

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 when data planning

Data planning should be part of evaluation planning processes (chapters 5) 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 former Department of Industry, Innovation and Science's Evaluation Ready approach highlighted data planning as an integral part of evaluation planning (DIIS 2017). The evaluation planning process involves developing a program logic, an evaluation strategy and a data matrix (box 9.7).

Data planning should consider how data are best collected, including who should collect the data, when they should be collected, and whether data collection methods are suitable for participants (box 9.8). 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 9.7 A data matrix to help with evaluation planning

A data matrix is a tool for organising evaluation questions, and identifying the information needed to answer 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 former 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 (2017); DJSB (2019).

Box 9.8 An example of adapting data collection processes for local contexts

Families and Schools Together 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 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.

New evaluation processes needed to be culturally relevant and meaningful to participants, while still being rigorous and consistent with the original, 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 include carefully considering what data are necessary for an evaluation, taking into account the costs of data collection relative to their value. 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).

Addressing sampling, non-response and attrition

Evaluations of mainstream policies and programs need to consider sampling needs, 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.

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.

Data collection, analysis and reporting also needs to 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.

One way to address the problems caused by having 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 9.9). Oversampling can improve the accuracy of results by increasing statistical power, and reducing statistical bias that can be problematic when working with small numbers.

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.

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. When planning for data collection, agencies should consider 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.

Box 9.9 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 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 their 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: BHI (2015, 2016).

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 (box 9.10). 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 (Mavromaras et al. 2018).

Aboriginal and Torres Strait Islander data governance and sovereignty

The quality and relevance of data used in evaluation can be improved by agencies and evaluators engaging better with Aboriginal and Torres Strait Islander people on data planning, collection and governance.

Participants spoke about limited engagement with Aboriginal and Torres Strait Islander people on what data are collected for evaluation. 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)

Box 9.10 National Disability Insurance Scheme in a remote location

As part of an evaluation of the National Disability Insurance Scheme (NDIS), this study looked at specific issues on the operation of the NDIS in a remote location — the Barkly region in the Northern Territory.

The study of the NDIS focused on seven main 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 that was more culturally sensitive and appropriate to remote needs and service delivery was recommended.

Source: Mavromaras et al. (2018).

Better engagement with Aboriginal and Torres Strait Islander people on decisions about data collection, management and use can:

- ensure that collected data reflect the priorities of Aboriginal and Torres Strait Islander people and communities
- improve trust in data collection and governance arrangements
- ensure that data are being collected, managed, used and stored in an ethical and culturally appropriate way
- build on existing Aboriginal and Torres Strait Islander data capability
- support Aboriginal and Torres Strait Islander data sovereignty.

The importance of engaging with Aboriginal and Torres Strait Islander people extends beyond evaluations, and there have been many calls by Aboriginal and Torres Strait Islander people for government to work better with Aboriginal and Torres Strait Islander people on data collection, management and use. The National Agreement committed governments to ‘shared access to data and information at a regional level’ with an objective that:

Aboriginal and Torres Strait Islander people have access to, and the capability to use, locally-relevant data and information to set and monitor the implementation of efforts to close the gap, their priorities and drive their own development. (JCOCTG 2020, pp. 3–4)

Further, parties to the National Agreement committed to:

... [put] partnerships in place between Aboriginal and Torres Strait Islander representatives and government organisations to guide the improved collections, access, management and use of data to inform shared decision-making for the benefit of Aboriginal and Torres Strait Islander people. (JCOCTG 2020, p. 14)

Many participants called for the Strategy to embed greater Aboriginal and Torres Strait Islander decision making about data collection, management and use through supporting Aboriginal and Torres Strait Islander data sovereignty⁵⁸. For example, the Lowitja Institute said:

We recommend that the Strategy and supporting Guide include principles of Indigenous Data Sovereignty. Many existing data and evaluation methods do not recognise or privilege Aboriginal and Torres Strait Islander knowledges or perspectives. (sub. DR157, p. 2)

And the CSIRO said:

Incorporating data sovereignty and intellectual property into the evaluation data model will strengthen the Strategy’s commitments to these rights. (sub. DR131, p. 3)

What is 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 (Maiam nayri Wingara Indigenous Data Sovereignty Collective and AIGI 2018a).

In June 2018, the Maiam nayri Wingara Indigenous Data Sovereignty Collective and the Australian Indigenous Governance Institute convened a National Indigenous Data Sovereignty Summit. The aim of the Summit was to progress Indigenous Data Sovereignty and Indigenous Data Governance through developing shared understandings and initiating

⁵⁸ See, for example: CSIRO (sub. DR131, p. 3); Kimberley Aboriginal Medical Services (sub. DR164, p. 2); Lowitja Institute (sub. DR157, p. 2); National Aboriginal and Torres Strait Islander Legal Services (sub. DR160, p. 2); Ninti One Ltd (sub. DR118, p. 4); Social Ventures Australia (sub. DR168, p. 1); UNSW Social Policy Research Centre (sub. DR127, p. 3); and Victorian Aboriginal Child Care Agency (sub. DR126, pp. 4-5).

Indigenous Data Governance protocols. Delegates attending the summit 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 (Maiam nayri Wingara Indigenous Data Sovereignty Collective and AIGI 2018a).

The CARE principles, designed by the Global Indigenous Data Alliance, articulate a shared understanding of Indigenous data sovereignty across the world. They are designed to sit alongside and complement already existing data standards. The CARE principles are:

- 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 (Research Data Alliance International Indigenous Data Sovereignty Group 2019).

Both sets of principles recognise the importance of Aboriginal and Torres Strait Islander data governance — that is, where Aboriginal and Torres Strait Islander people have the power to make decisions about data — in putting Aboriginal and Torres Strait Islander data sovereignty into practice. As Maiam nayri Wingara Indigenous Data Sovereignty Collective explained:

Indigenous data sovereignty is practiced through Indigenous data governance ... which asserts Indigenous interests in relation to data by:

- informing the when, how and why our data are gathered, analysed accessed and used; and
- ensuring Indigenous data reflects our priorities, values, culture, lifeworlds and diversity. (2018b, p. 1)

Aboriginal and Torres Strait Islander data sovereignty and evaluation

Participants noted the benefits of having Aboriginal and Torres Strait Islander data governance (box 9.11) for evaluations undertaken under the Strategy. The Commission's study on *Expenditure on Children in the Northern Territory* also highlighted the importance of community engagement on data and noted the value of community knowledge in contextualising and interpreting data. Several examples are cited in the report where community has been able to validate the 'on-the-ground realities behind data'. For example, the Palmerston bus services seemed sufficient, but community members pointed out that violence at the interchange discouraged the use of public transport which in turn resulted in inadequate public transport (PC 2020, pp. 209–210).

There are a number of ways evaluations can build in Aboriginal and Torres Strait Islander data governance (box 9.12). These include:

- engaging with Aboriginal and Torres Strait Islander people early on in an evaluation, including through 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 on or with the evaluation team to contextualise and interpret data.

Effectively implementing Aboriginal and Torres Strait Islander data governance requires flexibility, trust, commitment and technical capability. Bruhn (2014), examining several examples of Indigenous data governance from Australia, Canada and the United States, noted that:

- in negotiations of data sharing arrangements between governments and Indigenous organisations, trust-building and frequent contact between the parties were important in addressing the legal, policy, and capacity challenges that arise
- at the operational level, both commitment and technical capacity were required to implement data linkages
- some Indigenous data governance initiatives started small and built up over time, seeking support from senior levels as they progressed; others were initiated through high-level political accords that were then progressively implemented
- Indigenous data sharing arrangements varied in their interpretation and embodiment of Indigenous data governance principles. Indigenous partners supported approaches that allow for compromise, as long as they were assured that the spirit and intent of the principles were 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 is an important mechanism for ensuring that Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges are centred. In the context of mainstream policies and programs, the extent to which Indigenous data governance arrangements are built into evaluations needs to be balanced with the scope and size of the evaluation, and its impact on Aboriginal and Torres Strait Islander people.

Box 9.11 Participants point to the benefits of good data governance

National Aboriginal Community Controlled Health Organisation:

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)

The National Indigenous Australians Agency:

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)

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

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)

Sophia Couzos:

... 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 ... (sub. 92, p. 3)

Social Ventures Australia:

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)

Box 9.12 Examples of Aboriginal and Torres Strait Islander data governance in an evaluation

Stronger Communities for Children Program

Evaluation methods were developed with each 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.

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.

Indigenous Australians Health Program

This evaluation was designed with a steering group (more than half of the members were Aboriginal and Torres Strait Islander people). In the evaluation itself, 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.

Sources: Baily et al. (2018); Department of Health (sub. 35, attachment A); Just Reinvest NSW (sub. 80); KPMG (2018); Winangali Ipsos Consortium (2017).

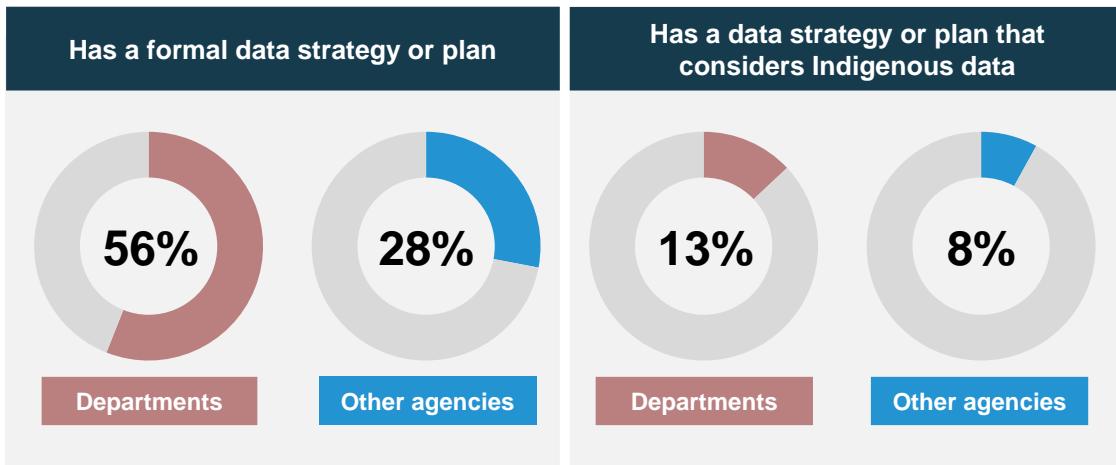
Aboriginal and Torres Strait Islander engagement in agency-level data governance

Aboriginal and Torres Strait Islander data governance should not just extend to data that are collected in evaluations. As outlined below, a key way to lower the costs and burdens of evaluations is by using existing data. Agencies should consider how to embed Aboriginal and Torres Strait Islander data governance in their general data planning activities to improve the quality and usefulness of data used for evaluation.

Some 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 9.2). This 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 9.2 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 52 Australian Government agencies: 16 departments and 36 other 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. Just 13 per cent of departments, and very few other agencies, reported considering Indigenous data as part of their general data planning activities, or having Indigenous data strategies or plans (figure 9.2). Some examples of agencies that do consider Aboriginal and Torres Strait Islander data in their data planning are provided in box 9.13. However, several agencies reported using the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) Guidelines for Ethical Research in Australian Indigenous Studies (2012) to guide how they collect, use and manage data about Aboriginal and Torres Strait Islander people.

Box 9.13 Aboriginal and Torres Strait Islander data planning in Australian Government agencies

Former Department of the Environment and Energy

The Department of the Environment and Energy's information strategy outlined key actions to improve the Department's approach to Aboriginal and Torres Strait Islander data. It committed 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.

Department of Health — Indigenous Health Division

The Indigenous Health Division's data strategy guides the use, improvement and development of data. It 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 to 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.

Indigenous Business Australia

Indigenous Business Australia has produced an impact framework to better assess how their policies and programs are generating impact. They have developed theories of change that map their programs back to the Interplay wellbeing framework. The framework identifies relevant indicators to identify economic, social and cultural impacts. These indicators sit alongside Indigenous Business Australia's existing performance monitoring and reporting systems, such as their corporate and annual reporting.

Sources: Department of the Environment and Energy (sub. 101, p. 5); Department of Health (pers. comm., 25 May 2020); IBA (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 consider:

- how Indigenous data governance can be strengthened
- what outcomes and indicators are important to Aboriginal and Torres Strait Islander people and whether they are available at a sufficiently disaggregated level to be useful
- 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.

Data gaps and the consistency of data collection across agencies

One of the challenges of collecting data that are 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. (2018a, 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 was developed over three years with 24 focus groups and two pilot ‘proof-of-concept’ studies (Jones et al. 2018).

The draft Strategy proposed a data dictionary of core indicators be developed that could be used across different evaluations. There was some support for the data dictionary.⁵⁹ The NSW Ombudsman, for example, said:

The Strategy’s suggestion for a ‘Data Dictionary’ is a welcome inclusion, as this has the potential for the related data to be representative of individual communities. It also has potential to represent community related data at a national level without losing meaningful understanding and direction. (sub. DR152, p. 2)

⁵⁹ See for example: Department of Health (sub. DR176, p. 5); Indigenous Data Network (sub. DR162, p. 2); and SNAICC (sub. DR152, p. 2).

Others suggested that the data dictionary should align with the National Agreement.

Recent negotiations on a revised Closing the Gap framework provides guidance on outcomes and indicators that are most informative for government and Aboriginal and Torres Strait Islander people. The outcomes and indicators include considerations of community, culture and connection to country, and have been the result of extensive consultation with Indigenous Australians. (NIAA, sub. DR161, p. 8)

... the highest priority indicators for inclusion in a data dictionary should be drawn from the new National Agreement on Closing the Gap. These indicators have been developed in partnership between governments and the Coalition of Aboriginal and Torres Strait Islander Peak Organisations, and so are already partially representative of the views of Aboriginal and Torres Strait Islander people. (Tasmanian Government, sub. DR166, p. 6)

However, others suggested that the data dictionary would overlap with initiatives under the National Agreement.

A comprehensive framework of terms, definitions, and principles for best-practice Indigenous data governance, together with an index of the location of data assets generated by and about, and relevant to, Indigenous people, held both by government and others, is already in the process of being developed. We submit that, rather than duplicating [the implementation of the priority reform on community data] by the [Australian Institute of Health and Welfare], [Indigenous Data Network] and Coalition of Peaks, government commit to supporting the current initiative. (Indigenous Data Network, sub. DR162, p. 3)

The National Agreement recognised that there are significant gaps in data about what matters to Aboriginal and Torres Strait Islander people. As the NIAA said:

The negotiations have highlighted that significant data development is needed if data is to be available to underpin an expanded, richer approach. This includes supporting local data collection efforts to provide an enhanced understanding of what is happening in different local contexts. (sub. DR161, p. 8)

As such, the National Agreement committed to a data development plan for the targets, drivers and contextual indicators that support each outcome of the National Agreement, including different dimensions of data disaggregation (JCOCTG 2020).

While the data development plan will be endorsed within two years of the National Agreement's commencement — and as such, detail on what it will entail is sparse — it appears the outcome indicator framework will be useful to identify where there may be data gaps. It takes into account the priorities of governments and Aboriginal and Torres Strait Islander people, and should provide a consistent and comparable set of indicators that map drivers and other contextual information back to outcomes.

Creating a data dictionary as part of implementing the Strategy therefore runs the risk of duplicating the work agreed to under the National Agreement. As such, we suggest that the data development plan and the fourth reform priority — which have both been committed to by governments — should be used as the basis for identifying and addressing data gaps. The

data development plan and community data initiative, as they are being developed, should also consider what data will be needed for evaluation.

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 to 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 can be costly. Wide coverage also allows for reliable comparison groups, who may be hard to contact or target in surveys, to be identified
- *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 to improve 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 for analysis. 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, chapter 3).

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 9.14). 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.

These reforms are a good first step to 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. The Independent Review of the Australian Public Service (APS) commented that:

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)

Box 9.14 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 released for consultation in September 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
- give advisory, guidance, regulatory and advocacy functions to the National Data Commissioner to oversee the scheme.

Sources: DPMC (2019c, 2020b); PC (2017a).

The National Agreement also calls for greater data sharing with Aboriginal and Torres Strait Islander people, communities and organisations. It commits government to share available, disaggregated regional data and information with Aboriginal and Torres Strait Islander organisations and communities. This is important for the Strategy for several reasons:

- it enables community-led evaluation efforts
- communities have access to data to make decisions about evaluations, such as what should be evaluated, what outcomes can be measured, and where further investigation may be required
- communities can use evaluation data to make decisions about changes to services in their community
- community knowledge can provide important context to the data that is shared with them.

Providing data back to evaluation participants is also part of ethical practice. As noted in the *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*:

Ownership, management and communication of research data and results should be negotiated between Indigenous peoples and the researcher or other parties based on the principles of Indigenous data sovereignty and governance. (AIATSIS 2020b, p. 19)

Agencies should engage with the community early about what data are needed for the community to effectively participate in evaluation planning, conduct and analysis as well as what data will be provided back to the community (and how it will be presented, for example, using infographics, comparisons and trends) at the end of an evaluation.

Data linkage

Data linkage can be a powerful tool to use in evaluations. It involves linking multiple datasets together by using features common of datasets (box 9.15). Data linkage can result in richer datasets that can be used to gain insights that are not possible with a single dataset or survey.

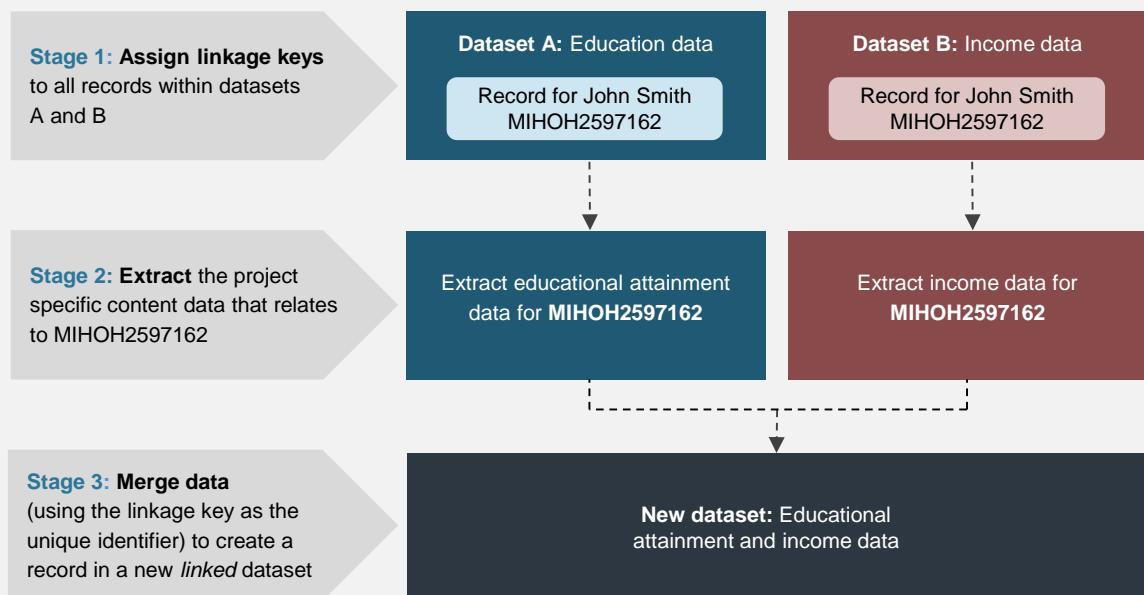
Box 9.15 The process of data linkage

Data linkage creates new datasets by combining existing data. The advantages include being able to make 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 for 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
- where Indigenous status has not been collected (or collected poorly) and Aboriginal and Torres Strait Islander people cannot be identified accurately in the data. Linking can be used to improve the quality of an Indigenous status variable
- 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.

However, the benefits of linked data need to be balanced with the ethical risks.

First, 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 2016, 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.

Second, 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 9.16), which will be adopted under the data sharing legislation and is also used by the ABS and the AIHW for their data sharing arrangements.

Box 9.16 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).

Last, the use of existing data that is not adequately contextualised risks presenting Aboriginal and Torres Strait Islander people only in terms of deficiency, rather than reflecting their strengths. As the NSW Aboriginal Housing Office noted:

... linked data can improve the quality and rigour of evaluation. However, the use of existing data, even when de-identified, is open to the risk of perpetuating deficit discourses about Aboriginal people as it is analysed and interpreted in the absence of other contributing external, contextual factors. (sub. DR130, p. 2)

As such, it is important that linked datasets include data that can provide the necessary contextual understanding.

The Australian, state and territory governments have invested significantly in integrated datasets that can be useful for evaluation. One example is the Data Integration Partnership for Australia (DIPA) (box 9.17). 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 ABS 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 9.17 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 Australian Government agencies involved in DIPA.

DIPA has several components. These include:

- *Data integration* delivered by the Australian Bureau of Statistics and the 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).

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 2019). 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.

Further improvements to DIPA datasets could improve their relevance for the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people. The National Agreement could be a framework to identify useful data linkages. It outlines both drivers and contextual information that are relevant for achieving outcomes. Linking these data together can give a more nuanced picture of what is needed to improve outcomes as well as being able to link the effects of policies and programs to outcomes.

DIPA was 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 Australian 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.

10 Monitoring and reviewing the Strategy

Key points

- Monitoring arrangements for the Indigenous Evaluation Strategy (the Strategy) have three important functions. These are to:
 - monitor the performance of Australian Government agencies against the Strategy
 - update government-wide evaluation priorities
 - monitor how effective the Strategy has been in encouraging high-quality and useful evaluations.
- These monitoring functions should be undertaken (in partnership) by the proposed Office of Indigenous Policy Evaluation (OIPE) and the proposed Indigenous Evaluation Council (the Council).
- The OIPE and the Council should produce regular, public reports that outline their findings and recommendations based on their 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:
 - assess the extent to which Australian Government agencies have complied with the Strategy and have implemented the Strategy effectively
 - identify good and/or innovative practice in the evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people
 - assess 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. The review should be led by an Aboriginal and/or Torres Strait Islander person and assess:
 - whether the Strategy's principles remain fit-for-purpose, and if not, what changes to the principles should be made
 - the extent to which the Strategy has been effective in encouraging higher quality and more useful evaluations, including the extent to which the key actions of the Strategy have been effective
 - the performance of those tasked with overseeing the Strategy and whether changes to the oversight arrangements are required.

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 how the Strategy should be monitored and reviewed.

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
- update government-wide evaluation priorities
- 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, design, reporting and use. If the Strategy is to improve the quality and usefulness of evaluations, agencies will need to follow this guidance. The Strategy also proposes a series of actions — most of which are for agencies — that need to be undertaken if the Strategy is to be implemented effectively.

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 is often done more for accountability or 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 will not change how agencies approach evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, or that only some agencies follow the guidance or implement the Strategy's actions, but not others.

External oversight means there will be a check on agencies to ensure that the evaluations they undertake are high-quality, are ethical, are useful and that their evaluation processes, practices and outputs align with the Strategy's principles, guidance and actions. As Ninti One said:

... the strategy and guide are all very well, but unless government agencies are held accountable for their commitment to evaluation and the findings that arise, then change will be limited.
(sub. DR118, p 2)

Or as Jobs Australia simply put it:

Compliance with the Strategy is central to its success. (sub. 57, p. 10)

However, while it is important to hold agencies to account for the way they implement the Strategy, this — in isolation — will not be sufficient to ensure the Strategy is successful. Growing a culture within the Australian Public Service (APS) where evaluation is valued, where it is undertaken consistently and capably and where findings are used to improve policies and programs affecting Aboriginal and Torres Strait Islander people will require more than just monitoring of the Strategy’s implementation. This is something that the Commission considered when designing the functions of the proposed Office of Indigenous Policy Evaluation (OIPE) (chapter 7).

In fact, if monitoring arrangements are enacted without support to help agencies improve their evaluation cultures and practices, there is a real risk that agencies will view the Strategy as little more than a compliance hurdle that needs to be overcome — or a box that needs to be ticked. If this were the case, the Strategy would have a limited impact on how policies and programs affecting Aboriginal and Torres Strait Islander people are evaluated, on how evaluations are used to improve these policies and programs, and ultimately, on the lives of Aboriginal and Torres Strait Islander people.

In addition to monitoring agencies’ compliance with the Strategy, monitoring arrangements should also be a platform where the use of good evaluation practices by agencies is 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 Lowitja Institute pointed out the benefit of such a function:

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. These lessons are also likely to be useful for those involved in evaluations beyond the APS, such as state and territory governments, evaluators and Aboriginal and Torres Strait Islander people, organisations and communities.

Updating government-wide evaluation priorities

The Commission has identified a set of ‘interim’ government-wide evaluation priorities to guide agencies’ efforts on what to evaluate under the Strategy (chapter 6). These interim priorities are based on:

- the outcomes of the National Agreement on Closing the Gap
- what Aboriginal and Torres Strait Islander people and communities told the Commission about what the priorities should be⁶⁰
- an initial assessment of the evidence base.

There is a need to formalise these government-wide interim priorities, and once formalised, to periodically review and revise them to ensure that they remain appropriate, reflect new information and centre Aboriginal and Torres Strait Islander people’s priorities about what should be evaluated. The monitoring process is an opportunity to do this.

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 relevant evaluation evidence. Monitoring arrangements should assess whether the Strategy encourages high-quality and useful evaluations in pursuit of that objective.

The Strategy (and the accompanying *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.

Because the Strategy is new and its implementation untested, as agencies begin to apply it, it may become apparent that some parts work well and some parts less so. 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.

General principles to underpin the design of monitoring arrangements

Monitoring arrangements should be independent ...

Independence must be enshrined in the monitoring arrangements for the Strategy if they are to be viewed as credible. Independence means that those doing the monitoring — free from the day-to-day operation of government or service delivery — can candidly examine and

⁶⁰ Feedback on evaluation priorities was received in submissions, a workshop on priority setting and in meetings.

report on 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, organisations and communities, government agencies, program managers and evaluators.

... transparent ...

To the greatest extent possible, the information feeding into — 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 effectively implementing the Strategy (else risk being publicly called out). As Maggie Walter said when commenting on the Australian National Audit Office's (ANAO's) report on the Indigenous Advancement Strategy:

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 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?

The Letter of Direction for this project points to a role for the Productivity Commission in reviewing agencies' conduct against the Strategy. When asking the Commission to develop the Strategy, the Treasurer said:

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. (Frydenberg 2019, p. 1)

Recognising that the successful implementation of the Strategy requires oversight functions that extend beyond just monitoring agencies' performance, the Commission is proposing the establishment of an OIPE. The OIPE would monitor agencies' conduct against the Strategy, but it would also have additional roles, including providing stewardship for the Strategy, championing the value of evaluation, and advising agencies on what the Strategy means in practice. The OIPE would undertake its monitoring functions in partnership with an Indigenous Evaluation Council (the Council) (chapter 7).

The OIPE could be housed in the Commission, although it could potentially also be housed in another government agency (chapter 7). However, assuming that the OIPE is responsible for monitoring the Strategy (as proposed), reflecting the principles outlined above, the agency where the OIPE is housed 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. Ideally, the agency would be a statutory authority, underpinned by legislation so that their independence and authority is formally established and enshrined. And because legislation can only be changed by the Parliament, a statutory authority has more stability, 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 (noting that the Council partnering with the OIPE would be one way to incorporate Indigenous governance into monitoring arrangements)
- 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 and credible 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.

In the longer term, the Commission is proposing 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 APS (chapter 7). The OIPE (including its monitoring functions) could be transferred to the CEE once it is established. This has several potential advantages, vis-à-vis retaining monitoring arrangements for the Strategy in an existing government agency.

- 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 it has all the tools 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, it would have more scope (and greater incentives) to gather and build evaluation expertise, 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 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.

How should monitoring occur?

The OIPE — in partnership with the Council — should produce regular reports that outline their findings and recommendations based on its monitoring of the Strategy's implementation. Having the OIPE and the Council jointly produce these reports is important, as both bodies will bring different strengths and perspectives to the monitoring process.

- The OIPE, with its role of overseeing the implementation of the Strategy within the Australian Government, would have in-depth knowledge of the Strategy's requirements and guidance, and the resources needed to undertake the 'day-to-day' monitoring of the Strategy's implementation.
- The Council, as a body that is made up of Aboriginal and Torres Strait Islander people, would bring an Indigenous lens to monitoring the Strategy's implementation and effectiveness. It would also bring lived experiences and expertise in the development, delivery and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people to complement the expertise of the OIPE.

In short, monitoring reports will be more credible, rigorous, complete and useful because of the combined input and expertise of the OIPE and the Council.

The monitoring reports should be publicly released — transparency will be crucial for improving 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. 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 against the Strategy be assessed?

In the first instance, the OIPE and the Council should examine agencies progress in implementing their basic responsibilities as set out in the Strategy. For example, the OIPE and the Council should assess whether departments (and agencies, if required to do so) have prepared and/or published a rolling Three Year Evaluation Forward Work Plan (as required by action 2 of the Strategy), and whether agencies have published evaluation reports (as required by action 7) and accompanying summaries and management responses (as required by actions 8 and 9).

The OIPE and the Council should also assess the degree to which agencies are upholding the principles of — and implementing the guidance provided by — the Strategy when evaluating policies and programs affecting Aboriginal and Torres Strait Islander people. In particular, this should involve examining how agencies are:

- determining what to evaluate and setting evaluation priorities
- planning, designing and conducting evaluations
- reporting evaluation findings
- building evaluation culture and capability.

The information needed to inform these assessments will need to come from a range of sources, including (but not limited to):

- departments' and agencies' rolling Three Year Evaluation Forward Work Plans. When examining these plans, the OIPE and the Council should consider the extent to which: evaluation priorities are determined based on an appropriate methodology; evaluation plans are clear; planned evaluations map to priority areas; and plans contain a mix of new and existing policies and programs to be evaluated
- the evaluations completed by agencies. When examining completed evaluations, the OIPE and the Council should consider the extent to which: evaluations match those identified in rolling Three Year Evaluation Forward Work Plans and high priority evaluations are taking place; the design, conduct and management of evaluations have been undertaken in a manner consistent with the Strategy and are high quality; and evaluations have been reported, translated and responded to effectively

-
- consultations with agencies about how they have planned for, conducted and used evaluations
 - views from Aboriginal and Torres Strait Islander people — as evaluators, evaluation participants and evaluation users — about the extent that agencies have effectively implemented the Strategy
 - views from other key stakeholders, such as evaluators and service providers.

The Commission has developed an initial ‘progression pathway’ as a tool for assessing the performance of agencies in implementing the Strategy. The progression pathway is an extension of the ‘maturity approach’ flagged in the draft Strategy. Several participants supported a maturity approach and/or encouraged the Commission to further articulate what different levels of maturity look like in the context of the Strategy’s implementation.⁶¹

The progression pathway identifies what unsatisfactory, developing, mature and leading practice looks like for each of the Strategy’s principles for the various stages of an evaluation. The pathway was informed by other evaluation maturity matrices (such as those of the former Department of Industry, Innovation and Science (DIIS 2017, pp. 38–39) and the National Indigenous Australians Agency (published in ANAO (2019b, pp. 53–55)), as well as the views of project participants about what characterises poor, good and leading practice given the principles and actions of the Strategy. The progression pathway is available on the Commission’s website.

Using a progression pathway to inform assessments of agencies’ performance recognises that agencies will be starting from different bases, and as such, it may take some agencies time to change their evaluation culture and develop good evaluation practice. The pathway provides clear and transparent benchmarks to inform assessments of agencies’ conduct, and for agencies to look to as they seek to improve their evaluation practices.

Through their reporting, the OIPE and the Council should identify agencies who have not met their responsibilities under the Strategy, or have implemented the Strategy in an unsatisfactory manner. However, in order for the monitoring process to have the greatest effect on raising the quality of evaluations, it needs to do more than identify instances of non-compliance. It should also provide agencies with the information they need to improve their practices, including by:

- identifying and highlighting examples of good practice — such as rolling Three Year Evaluation Forward Work Plans or individual evaluations that are exemplary — and/or providing case studies detailing how good practice can be undertaken

⁶¹ Participants included: BetterEvaluation (sub. DR144, p. 3); National Voice for our Children (SNAICC) (sub. DR163, p. 7); and the Department of Health (sub. DR176, p. 2). The Independent Members of the National Indigenous Australians Agency Evaluation Committee also suggested that rubrics could be developed that outline what poor, acceptable and good practice looks like for each of the Strategy’s principles (sub. DR115, p. 3).

-
- determining and disseminating lessons about what approaches, methodologies and practices are or are not working well when evaluating policies and programs affecting Aboriginal and Torres Strait Islander people, based on the evaluations completed by agencies
 - providing examples where following the Strategy has resulted in better evaluations, and consequently, improvements to policies and programs that affect Aboriginal and Torres Strait Islander people
 - providing recommendations to agencies on how they can improve their evaluation planning, commissioning, conduct or reporting, or how they can improve their evaluation culture and capability.

How should government-wide evaluation priorities be updated?

As part of the reporting process, the OIPE and the Council should assess whether the existing — and in the first instance, the interim — government-wide evaluation priorities remain appropriate given:

- changes in the nature and mix of policies and programs delivered by Australian Government agencies that affect Aboriginal and Torres Strait Islander people, including those to implement policy and reform priorities identified in the National Agreement on Closing the Gap
- changes in the evidence base around ‘what works’ for policies and programs affecting Aboriginal and Torres Strait Islander people, recognising that evidence ‘gaps’ may emerge, close or change as agencies undertake better quality evaluation
- the evaluation priorities of Aboriginal and Torres Strait Islander people, recognising that the Commission is proposing that priorities be set through a joint partnership arrangement with Aboriginal and Torres Strait Islander people
- evaluations that agencies propose to undertake in the future, as identified in rolling Three Year Evaluation Forward Work Plans
- the views of agencies and other key stakeholders on what evaluation priorities should be.

If, when undertaking this assessment, the OIPE and the Council consider that there would be benefit in changing the existing government-wide evaluation priorities, they should provide recommendations to the Australian Government on what the government-wide evaluation priorities should be.

The OIPE and the Council should also consider where there might be opportunities for agencies to collaborate to undertake evaluation, given government-wide evaluation priorities. This will require the OIPE and the Council to examine rolling Three Year Evaluation Forward Work Plans to identify areas of commonality, and where common evaluation areas are identified, to assess whether there are likely to be benefits from agencies collaborating in these evaluations. Such benefits may include — for example — greater evaluation coverage, greater scope to examine interactions between policies and programs, or resource savings arising from collaboration.

How should the Strategy's effectiveness be monitored?

Monitoring the Strategy's effectiveness should involve examining the extent to which the Strategy (and the accompanying *Guide to Evaluation under the Indigenous Evaluation Strategy*) has improved the quality and usefulness of evaluations. It should also include looking at whether there is scope to improve the Strategy's implementation and/or to update or improve its content or the accompanying guidance material and progression pathway. The Strategy should be a living document that is updated to reflect changes in what is considered to be good practice. Any issues in the way the Strategy is working that could reduce its effectiveness should also be addressed (this will be especially important in the first few years because the Strategy will be new and its implementation untested).

Monitoring the Strategy's effectiveness should involve assessing the extent to which:

- the Strategy's priority setting process is directing evaluation effort to high priority areas
- the Strategy is resulting in Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges being centred in relevant evaluations
- the Strategy, and its guidance material and progression pathway, reflect 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.

Information to assess the Strategy's effectiveness will come from the OIPE's and the Council's monitoring of agencies' performance against the Strategy. However, insights from the experiences of those overseeing, undertaking, participating in, using or impacted by evaluations, will be important for reaching conclusions about the Strategy's effectiveness. For this reason, it is important that there are ways people can share their views on how the Strategy is working. This could include a public submission process, targeted consultations, roundtables and/or workshops.

Recommendations could 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 would 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, there may not be a strong enough incentive for agencies to start changing how they evaluate policies and programs affecting Aboriginal and Torres Strait Islander people (in effect, ‘putting it off until later’ knowing that an assessment of their performance will not occur for some time). The timing of the first monitoring report for the Strategy needs to balance these 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 superficial level. This would be far from ideal, given much of the value from the proposed monitoring process will come 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 too regular, and demand for consultation 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 consider that the first monitoring report should be provided to the Australian Government (and released publicly) two years after the Strategy is first endorsed.

This will 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 the quality of evaluations undertaken under the Strategy in detail — there may be too few completed evaluations to allow for informed judgements — it should be more than enough time to assess agencies' performance in incorporating the Strategy’s principles and guidance into their evaluation planning (particularly because departments will be expected to have an initial rolling Three Year Evaluation Forward Work Plan within a year of the Strategy being endorsed (chapter 6)).

After the release of the first report, monitoring reports should be produced biennially. While the first report would be expected to focus predominately on evaluation planning, subsequent reports should 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?

It is also important to establish a separate process for reviewing the Strategy itself. Such a review process should assess:

- whether the Strategy’s principles remain fit-for-purpose
- the extent to which the Strategy has been effective in encouraging higher quality and more useful evaluations (this includes assessing the effectiveness of the Strategy’s actions)
- the extent to which oversight arrangements remain fit-for-purpose, including by assessing the performance of those overseeing the Strategy and whether changes to the oversight arrangements are required.

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. The review process should involve broader and deeper assessments and analysis of the Strategy’s effectiveness than those of the monitoring process, and the resources 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. They seek to influence all aspects of evaluation — including priority setting, evaluation planning, evaluation conduct, commissioning practices and evaluation reporting — as well as how knowledge is shared and evaluation culture and capability is developed. The principles also form the basis of the progression pathway, which maps a vision of what the successful implementation of the Strategy looks like. Given the principles underpin so many elements of the Strategy, it is important that they remain fit-for-purpose and useful if the Strategy is to be most effective.

While the monitoring process is well-placed to assess whether the Strategy’s principles are being upheld by agencies in their evaluation practices (or whether more guidance on their application is required), 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 to which 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 systematic process to determine evaluation priorities; developing and publishing rolling Three Year Evaluation Forward Work Plans (if required to do so); preparing Aboriginal and Torres Strait Islander Impact Assessment and Evaluation Plans for new policies and programs; and publishing evaluation reports with a timely management response.

The monitoring process will assess agencies' progress and performance implementing these actions, and provide recommendations on how their implementation can be improved. But the monitoring process will not ask 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

It is also essential that the oversight arrangements for the Strategy work well. While the Strategy ultimately places the responsibility on agencies to undertake better evaluation of policies and programs that affect Aboriginal and Torres Strait Islander people, the oversight arrangements (chapter 7) help to provide the support, tools and the incentives agencies need to implement their responsibilities effectively.

A review process provides an opportunity to examine the effectiveness of the oversight arrangements and whether there is scope to improve them. A review of the Strategy should consider:

- *the effectiveness of the OIPE and the Council.* Reviewing the performance of the OIPE and the Council is valuable for two reasons. First, external review provides an incentive for the OIPE and the Council to invest the resources needed to undertake their functions effectively. Second, it is an opportunity to identify improvements to the way the OIPE

and the Council go about their work, which in turn could make the Strategy more effective

- *whether there have been changes to the policy environment that justify changes to the oversight arrangements for the Strategy.* The policy landscape in which the Strategy will operate is dynamic and subject to considerable change (chapters 1 and 2). For example, the Strategy will be implemented at the same time as Australian Government agencies (along with other jurisdictions) make changes to the way they work with Aboriginal and Torres Strait Islander people because of their commitments under the National Agreement on Closing the Gap. The Australian Government has also agreed to a series of actions related to evaluation as part of its response to the Independent Review of 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 its oversight arrangements — are working. This suggests a reasonable period of time 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’, particularly given the policy environment around the Strategy is changing. It is important that the review process is able to consider how the Strategy best interfaces with these changes.

In the draft Strategy, the Commission proposed that the Strategy should be reviewed five years after it is first endorsed. Some participants suggested that this was too far into the future or that there should be an interim review of the Strategy. The Department of Health (sub. DR176, p. 7) suggested that an interim review should occur two years after the Strategy is first implemented. The Queensland Nurses and Midwives Union (sub. DR128, p. 4) also suggested an interim review process.

Given the proposed scope of the review process — as set out above — there is a significant risk that any review process that commences within five years of the Strategy’s endorsement will not have the information needed to make valid conclusions about how the Strategy and its oversight arrangements are functioning, and how they can be improved. For example, if a review of the Strategy was to commence two years after the Strategy is endorsed, this would mean that the review would potentially be completed very soon after the first monitoring report is released. A review five years after the Strategy is endorsed would mean that information from two reporting periods could feed into the review, allowing for more informed judgements about the Strategy’s effectiveness and the ongoing suitability of its oversight arrangements.

That said, the OIPE and the Council would monitor the extent that the Strategy has been effective in encouraging high-quality and useful evaluation, which would include assessing whether there are opportunities to update the Strategy or to improve its implementation. Where such opportunities are identified, the OIPE and the Council could make recommendations to agencies — or in the event that changes at a whole-of-government level would facilitate better implementation, the Australian Government — to action these. This means that ongoing opportunities to revise the Strategy — should this result in its improved implementation — are built into the monitoring and reporting process.

Some important features of the review process

Given we are proposing the review 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
- *the review being led by an Aboriginal and/or Torres Strait Islander person* — the Strategy's objective is to improve the lives of Aboriginal and Torres Strait Islander people by having policies and programs informed by high-quality and useful evidence. A review that is led by an Aboriginal or Torres Strait Islander person is likely to be best placed to assess the extent that the Strategy has been successful in pursuit of this aim. It may also add to the credibility of the review process.

11 Bringing it all together: an Indigenous Evaluation Strategy

Key points

- The Indigenous Evaluation Strategy (the 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
 - 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 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 — should be interpreted.
- The Strategy would be strengthened by the Australian Government endorsing it and requiring Australian Government agencies to comply with it.

This chapter brings together insights from the earlier chapters of this background paper and looks at what should be included in the Indigenous Evaluation Strategy (the Strategy). As discussed in chapter 1, the Strategy was not co-designed with Aboriginal and Torres Strait Islander people as some participants suggested it should be. However, the Commission drew extensively on the views of participants, including those expressed in meetings, roundtables, workshops and submissions.

The key components of the Strategy are outlined in section 11.1. The principles for the Strategy are discussed in section 11.2.

11.1 The key components of the Strategy

The main purpose of the Strategy is to provide 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.

The Commission was asked to establish a principles-based framework, identify priorities for evaluation, and set out an approach to reviewing agencies' conduct of evaluations against the Strategy (chapter 1). While these are three key components of any 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 for an effective Strategy. The Strategy also needs a clearly defined objective, 'actions' to support a culture of evaluation across the Australian Public Service, and set out who is responsible for what under the Strategy.

As the Australian National Audit Office's review of the Indigenous Advancement Strategy Evaluation Framework pointed out:

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 key components of the Strategy 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 list of interim evaluation priorities and a process for reviewing and setting ongoing 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).

The components are 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 to identify what success will look like. As discussed in chapter 1, there is widespread agreement that the objective of the Strategy should be to improve the lives of Aboriginal and Torres Strait Islander people. The Strategy also needs to set out how this objective will be achieved, noting that it is better policies and programs, not evaluation per se, that will improve the lives of Aboriginal and Torres Strait Islander people.

A principles-based framework

Using 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 2013c) and the American Evaluation Association's *Guiding Principles* (AEA 2018).⁶² Principles-based frameworks for evaluation are also used in other Australian, state and territory government evaluation strategies (chapter 2).

Participants supported a principles-based framework to guide evaluation practice under the Strategy (box 11.1). One of the identified advantages of a principles-based approach is that it provides 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. Another was that a principles-based approach would allow the Strategy to adjust to a changing policy landscape.

⁶² The Australasian Evaluation Society is now known as the Australian Evaluation Society.

Box 11.1 Participants recognised the advantages of a principles-based framework

Chelsea Bond et al.:

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 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:

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)

Institute for Human Security and Social Change:

There is much to be commended in the Strategy, in particular the principles-based approach which provides for a high degree of adaptation and flexibility in evaluation approach and implementation. (sub. DR119, p. 1)

Empowered Communities:

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)

Ernst & Young:

The [Indigenous Evaluation Strategy] presents a principles-based approach which is welcomed by [Ernst & Young], particularly in the context of the evolving relationship between First Nations’ communities and Government, with genuine partnership its core tenet. (sub. DR140, p. 3)

And 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.

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. (Department of Health, sub. 35, p. 1)

Practical resources may also need to be made available to support Government agencies to translate the evaluation framework into practice. (Western Australian Government, sub. 74, p. 2)

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. (Maggie Walter, 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 lifting 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 how to build capability and a culture of evaluation.

Balancing the need for the Strategy to be a concise document while also providing guidance, the Strategy has a companion document — *A Guide to Evaluation under the Indigenous Evaluation Strategy*, which provides additional guidance to agencies on implementing the principles.

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 appropriate governance arrangements. These need to clearly define who is responsible for what and embed incentives to change behaviour.

While primary responsibility for conducting evaluations under the Strategy should continue to be with Australian Government agencies, centralised leadership, support, coordination and oversight is needed if the Strategy is to make a difference to evaluation processes and practice. The Strategy outlines the proposed governance arrangements for the Strategy, including how such arrangements ensure effective engagement with Aboriginal and Torres Strait Islander people (chapter 7).

A process for setting evaluation priorities

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 the 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)

As discussed in chapter 6, there was strong support from participants for government-wide evaluation priorities to be aligned with the National Agreement on Closing the Gap (the National Agreement).

The Strategy outlines a process for setting evaluation priorities — both at the government-wide and agency level. Because it will take some time for the Strategy’s governance arrangements to be put in place, the Strategy identifies interim priority areas — based on the National Agreement — to guide agencies’ efforts on what to evaluate under the Strategy. These priority areas should be formalised by the proposed Office of Indigenous Policy Evaluation in partnership with the proposed Indigenous Evaluation Council (chapter 7).

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. External oversight of agencies’ conduct of evaluations against the Strategy will provide an incentive for agencies to ensure 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 (chapter 10).

Monitoring under the Strategy will be more than checking agencies’ compliance with the Strategy. It will also focus on highlighting good evaluation practices and sharing lessons on what works (or does not work) 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)

Monitoring should also include an assessment of whether the Strategy itself encourages high-quality and useful evaluations. 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 is fit-for-purpose and useful (chapter 10).

There should also be an independent review of the Strategy undertaken five years after it is endorsed to assess: whether the Strategy’s principles remain fit-for-purpose; the performance of the 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 these important roles.

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 5, 6, 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 a number of 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 said:

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)

The Strategy would be strengthened by the Australian Government endorsing it and requiring Australian Government agencies to comply with it.

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'.

There is a high degree of consistency in the principles 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 Indigenous Advancement Strategy Evaluation Framework (DPMC 2018b) and Indigenous Business Australia's Impact Framework (2020) are notable exceptions.

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; self-determination; 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).

Others advocated principles from existing frameworks to inform the Strategy, including:

- the Lowitja Institute's *Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health* (Kelaher et al. 2018b)⁶³
- the BetterEvaluation *Ethical Protocol for Aboriginal and Torres Strait Islander Evaluation* (Gibb et al. 2019).⁶⁴

Principles for the Strategy

The principles for the Strategy have two purposes:

1. to guide agencies and evaluators when they are planning, conducting, reporting and using evaluation
2. to guide priority setting, knowledge sharing and translation, capability building , monitoring and review under the Strategy.

⁶³ Suggested by National Aboriginal Community Controlled Health Organisation (sub. 95, p. 4), Indigenous Allied Health Australia (sub. 31, p. 4), Australian Evaluation Society (sub. 49, p. 23), Aboriginal Health Council of South Australia (sub. 61, p. 6), Close the Gap Campaign Steering Committee (sub. 71, p. 4) and Aboriginal Medical Services Alliance Northern Territory (sub. 81, p. 2).

⁶⁴ Suggested by National Aboriginal Community Controlled Health Organisation (sub. 95, p. 4), Aboriginal Health Council of South Australia (sub. 61, p. 6) and Fred Hollows Foundation (sub. 14, p. 11).

Table 11.1 Principles for the Strategy suggested by participants

<i>Victorian Aboriginal Community Controlled Health Organisation</i>	
<ul style="list-style-type: none"> • Recognition • Self-determination • Community building: do-no-harm and strength-based approaches • Ethics 	<ul style="list-style-type: none"> • Accountability • Transparency • Effectiveness • Cultural continuity
<i>Victorian Aboriginal Child Care Agency</i>	
<ul style="list-style-type: none"> • Self-determination • Culturally informed approaches • Accountability • Strengths-based approach and one that is centred on the aspirations of Aboriginal people 	<ul style="list-style-type: none"> • Aboriginal-led collaboration and partnership • Person-centred • Community engagement • Equity
<i>Aboriginal Health Council of Western Australia</i>	
<ul style="list-style-type: none"> • 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 	<ul style="list-style-type: none"> • 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>	
<ul style="list-style-type: none"> • Respectful • Responsive • Reciprocal • Competent 	<ul style="list-style-type: none"> • Accountable • Transparent • Ethical • Integrity
<i>Australian Evaluation Society</i>	
<ul style="list-style-type: none"> • 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 	<ul style="list-style-type: none"> • 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).

From the long list of possible principles, the Commission proposes 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.

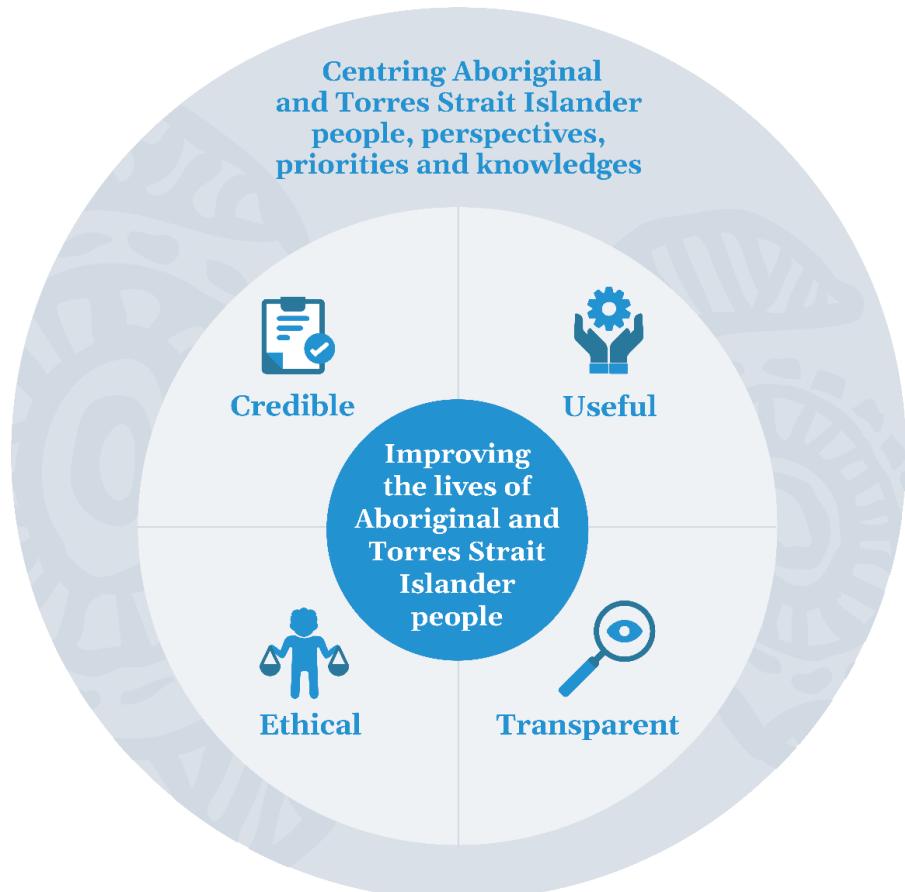
The Commission proposes that the overarching principle of the Strategy be **centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges** and that the overarching principle is the lens through which the Strategy’s other proposed principles — **credible, useful, ethical and transparent** — should be interpreted (figure 11.2).

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 the policies and programs have limited value.

The proposed overarching principle is about recognising the strengths and diversity of Aboriginal and Torres Strait Islander people, communities, knowledges, histories and cultures. It is about building 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.

Figure 11.2 Guiding principles of the Indigenous Evaluation Strategy



The idea that Aboriginal and Torres Strait Islander people should be at the centre of the Strategy supports calls by many participants for the principles to include: support for self-determination, data sovereignty and empowerment; partnership, engagement and co-design; recognition of Aboriginal and Torres Strait Islander cultures, knowledges and methodologies; and drawing on the perspectives, priorities and knowledges of Aboriginal and Torres Strait Islander people (table 5.3, table 11.1, box 11.2). It also aligns with the National Agreement's commitments for governments to work differently with Aboriginal and Torres Strait Islander people through partnerships and shared decision making.

Box 11.2 Support for the overarching principle

Lowitja Institute:

The Lowitja Institute strongly supports the core principle of the strategy to centre Aboriginal and Torres Strait Islander peoples, knowledges, perspectives and priorities. Evaluation needs to align with the aspirations of Aboriginal and Torres Strait Islander peoples so that it can lead to better designed policies and programs that serve to benefit Aboriginal and Torres Strait Islander peoples. (sub. DR157, pp. 1–2)

Australian Indigenous Doctors' Association:

AIDA is especially supportive of the overarching principle of 'centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges' being used as the lens through which to view and interpret the Strategy's other principles – credible, useful, ethical and transparent. AIDA agrees that by embedding evaluation processes viewed through an Indigenous lens at every stage of government project/program development, delivery and implementation, those projects/programs are more likely to have beneficial outcomes for Aboriginal and Torres Strait Islander peoples. (sub. DR125, p. 1)

Australian Medical Association:

The AMA supports placing the perspectives, priorities, and knowledge of Aboriginal and Torres Strait Islander people at the centre of the Strategy, and it must be ensured that all Commonwealth departments are bound by this principle in practice. The AMA acknowledges that the perspectives of Aboriginal and Torres Strait Islander people vary across the country and that evaluations of policies and programs will differ between program areas and locations. This is something that must be recognised by Commonwealth departments when conducting evaluations, and it must be acknowledged that these differences are beneficial to evaluations. (sub. DR116, p. 2)

Wuchopperen Health Service:

The overarching principle of the Strategy in 'centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledge' is integral to the strength of the strategy and its successful implementation and is endorsed by Wuchopperen. (sub. DR117, p. 1)

Ninti One:

The efficacy of the [Indigenous Evaluation Strategy] is bound to the principle of centring Aboriginal and Torres Strait Islander people perspectives, priorities and knowledges. However, we suggest that it ought to be clearly acknowledged within the documents that there is no one single Indigenous perspective, and that these perspectives, as well as priorities can vary within locations. A diversity of Indigenous voices, at all stages, including interpretation of the results, will add robustness and rigour. (sub. DR118, p. 2)

Institute for Human Security and Social Change:

Giving priority to the overarching principle of centring Aboriginal and Torres Strait Islander people's perspectives, priorities and knowledge is also to be commended and we believe, if correctly implemented, could support the Strategy to make a substantial contribution to meaningful assessment of Australian government programs and policies. (sub. DR119, p. 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 'Objectives, principles and defining success' roundtable.

The overarching principle of Strategy is strongly supported (box 11.2).⁶⁵ However, some called for more explicit acknowledgement of self-determination and data sovereignty in the

⁶⁵ Participants were also largely supportive of the other proposed principles, although a small number suggested additional principles, including building capacity (Victorian Aboriginal Child Care Agency, sub. DR126, p. 2), cultural responsiveness (Yulang Indigenous Evaluation, sub. DR150, p. 13), Indigenous

discussion about what the centring principle meant in practice (box 11.3). Others argued that the overarching principle did not sufficiently emphasise the potential role of Aboriginal and Torres Strait Islander people in leading evaluations. For example, the Victorian Aboriginal Child Care Agency said it:

... welcomes the overarching principle of the [Indigenous Evaluation Strategy] Draft of centring Aboriginal and Torres Strait Islander people and the way that engagement and partnership are encouraged in the policy and draft. However, the draft does not demonstrate how this will result in Aboriginal leadership of evaluation. ... Greater clarity and some additional emphasis on leadership will ensure that the draft supports a shift from Aboriginal and Torres Strait Islander people, Communities and organisations having evaluation *done to* them to being *done with* them through genuine partnerships. (sub. DR126, p. 3)

In light of this feedback, further explanation has been added to the Strategy on how the centring principle supports self-determination and data sovereignty, and the role of Aboriginal and Torres Strait Islander people as partners and leaders in evaluation.

As a key point of difference to other Australian Government agency evaluation frameworks, evaluators and agencies are likely to require more guidance on implementing this principle than the other principles. As the Institute for Human Security and Social Change said:

It is the experience of the Institute that centring Aboriginal and Torres Strait Islander people's perspectives, priorities and knowledges in the evaluation process will raise considerable challenges to traditional evaluation practice. The worldviews and values which are encountered in consultation with Aboriginal and Torres Strait Islander people are often different to the values and worldviews of government and other white Australian stakeholders, creating inherent tensions and challenges for evaluation framing and implementation. This has implications for the other principles outlined in the Strategy together with the ideas included in the Strategy guide. (sub. DR119, p. 2)

Credible

This principle is about ensuring that evaluation processes, methods, 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.

Credibility is also about the processes used during evaluation planning and conduct to verify the results. Credibility 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).

Data Sovereignty (Lowitja Institute, sub. DR157, p. 2) and healing and social and emotional wellbeing (Australian Evaluation Society, sub. DR174, p. 5).

Box 11.3 Calls for self-determination and data sovereignty to be referred to in the Strategy

National Aboriginal and Torres Strait Islander Legal Services:

... we consider that explicit mention of our self-determination in the Strategy itself is critical, particularly as Australia is a signatory to the *United Nations Declaration on the Rights of Indigenous Peoples* as well as other human rights international legal instruments that require Australia to advance the self-determination of our people. We note that inclusion of self-determination as an express objective in the Strategy will also contribute toward ensuring our community's engagement with the Strategy. (sub. DR160, p. 1)

Community First Development:

While we broadly support the draft Strategy's overarching principle ... and other guiding principles ... we would like to reiterate that evaluation in First Nations contexts must be consistent with the Australian Government's endorsement of the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP), First Nations peoples' right to self-determination, and the four Priority Reforms outlined in the new National Agreement on Closing the Gap ... (sub. DR134, p. 4)

Queensland Family and Child Commission:

The QFCC is concerned by the absence of any direct reference to the rights of Aboriginal and Torres Strait Islander peoples and their self-determination in the Strategy. The Strategy instead uses terminology such as 'centring' and 'engaging' Aboriginal and Torres Strait Islander peoples, without explicitly stating how this is embedded in the Strategy, or how the Strategy can ensure active efforts to embed these concepts in future policy and programs. ... The current language builds an overall impression that the Strategy is 'done to' rather than 'by' Aboriginal and Torres Strait Islander peoples for Aboriginal and Torres Strait Islander peoples. (sub. DR136, p. 4)

Empowered Communities:

The first principle of 'Centring Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledge' is positive but not sufficient, and should include empowerment and Indigenous agency as core to the 'centring' principle. (sub. DR171, p. 4)

Fred Hollows Foundation:

The Strategy should acknowledge and explicitly address this imbalance of power, through a commitment to upholding the right to self-determination and free, prior and informed consent, and to supporting Indigenous data sovereignty. ... Indigenous data sovereignty and self-determination are reliant on each other, and without a commitment to both, the Strategy will not meet its core objective, or uphold its guiding principle. (sub. DR156, pp. 1, 3)

Victorian Aboriginal Child Care Agency:

While there is reference in the draft strategy to developing protocols around data and the need for partnerships with Aboriginal and Torres Strait Islander people in the governance, development, collection, use and management of data, there is no reference to data sovereignty in the draft strategy or guide. (sub. DR126, pp. 4–5)

Vision 2020 Australia:

... Vision 2020 Australia recommends the Draft Strategy actively acknowledge and commit to Indigenous Data Sovereignty. This would reinforce the right of Aboriginal and Torres Strait Islander people to exercise ownership and governance over Indigenous data through the creation, collection, access, analysis, interpretation, management, dissemination and reuse of that data to inform the design and implementation of locally determined solutions. (sub. DR135, p. 4)

In line with the overarching principle, the Strategy should consider credibility from the perspective of Aboriginal and Torres Strait Islander people. This emphasises the importance

of drawing on Aboriginal and Torres Strait Islander perspectives and knowledges, and using culturally appropriate methods. As Yulang Indigenous Evaluation said:

Aboriginal and Torres Strait Islander people are more likely to accept an evaluation as credible if it is carried out in a culturally responsive manner, with the issues that are important to Aboriginal and Torres Strait Islander people being explored in a way that Aboriginal and Torres Strait Islander people see as relevant. (sub. DR150, p. 14)

The Fred Hollows Foundation also argued that credibility can be enhanced by:

... engaging local evaluators who understand the context and have credibility with their communities. This will enhance the credibility of Australian Government evaluations and help address power imbalances. (sub. DR156, p. 3)

And the CSIRO said:

... for some programs or policies, better and more credible evaluations may be achieved if led or co-designed by Aboriginal and Torres Strait Islander evaluators and/or organisations. (sub. DR131, p. 8)

Credible evaluations require evaluators who are technically and culturally capable, and 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 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 to adopt approaches and methods that are rigorous, fit-for-purpose and answer the questions that policymakers and Aboriginal and Torres Strait Islander people want 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)

Useful

This principle is about evaluations that are useful for decision makers and Aboriginal and Torres Strait Islander people alike. 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. Evaluations that do not provide useful results are a waste of resources, a point made by many participants (box 11.4).

Evaluations are most useful when they are of high quality and are timed to contribute to policy-making decisions. This means the right questions are asked, evaluators have engaged with those affected by the policy or program, and evaluations are rigorous and impartial.

Useful evaluations consider the needs of evaluation users in the planning phase (chapter 7). This includes understanding what is important from the viewpoint of Aboriginal and Torres Strait Islander people, organisations and communities, as well as policy makers in terms of what to evaluate and what outcomes to measure as well as identifying where there are gaps in the evidence base.

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. This will mean that evaluation findings will be available at key decision points during implementation, and support continuous quality improvement. On this point, the NSW Aboriginal Land Council, 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)

The usefulness of individual evaluations can also be enhanced by synthesising evaluation findings into the existing evidence base and using meta-analysis of multiple evaluations to provide broader insights into policy and program effectiveness and evaluation quality.

Box 11.4 Participants' comments on the importance of useful evaluations

Australian Evaluation Society:

Irrespective of the framing, design, quality and conduct of Indigenous evaluation or evaluation with Indigenous peoples; the evaluations come to nothing if the findings do not consistently inform Indigenous policy and planning processes — locally, regionally and nationally. (sub. DR174, p. 3)

Empowered Communities:

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)

The Smith Family:

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)

The Fred Hollows Foundation:

A commitment to consider an evaluation's 'usefulness' from both an Australian Government perspective and from the perspective of Aboriginal and Torres Strait Islander communities. This means that Aboriginal and Torres Strait Islander Peoples involved in or affected by an evaluation should benefit from it. (sub. DR156, p. 4)

Victorian Aboriginal Child Care Agency:

... 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)

The Tasmanian Government:

... due to time and other resourcing constraints there is often a temptation when conducting evaluations to construct the evaluation questions around what data is already available, rather than considering what questions and data might provide the most useful findings. (sub. DR166, p. 4)

Close the Gap Campaign Steering Committee:

... useful results that are not used or incorporated into a process of continuous process and quality improvement and embedded into a policy cycle are an additional waste of resources and undermine confidence in the entire process of evaluation. We support the view in the draft that evaluations should be embedded in the policy cycle ... (sub. DR159, p. 2)

Ethical

While it can be argued that ethical conduct is essential for evaluations to be both credible and useful, the Commission proposes '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)

A number of participants pointed out the benefits of ethical evaluation practices for improving research quality and improving the experiences of Aboriginal and Torres Strait Islander people involved in research and evaluation:

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. (Indigenous Community Volunteers, sub. 88, p. 18)

Seeking approval from ethics committees is a valuable part of the research process as it ensures research plans adhere to ethical guidelines, which is especially important for Aboriginal and Torres Strait Islander peoples who have had particularly negative experiences of research. (UNSW Social Policy Research Centre, sub. DR127, p. 6)

VACCA appreciates the importance of stressing the need for ethical design and implementation even where [Human Research Ethics Committee] approval is not required. (VACCA, sub. DR126, p. 3)

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 the:

- *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research* (AIATSIS 2020b)
- *Ethical Conduct in Research with Aboriginal and Torres Strait Islander Peoples and Communities* (NHMRC 2018a)
- *National Statement on Ethical Conduct in Human Research* (NHMRC, ARC and UA 2018).

Rather than being definitive about when formal ethical review is required during evaluation, the Commission proposes that evaluations 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)

A more systematic approach to ethics in evaluation is likely to increase the cost and time to undertake evaluation. However, the benefits — that is, higher quality evaluations that Aboriginal and Torres Strait Islander people can have confidence are conducted in line with best practice ethical principles — are expected to outweigh the costs. However, it will be important to ensure that agencies are not avoiding evaluations of policies and programs that affect Aboriginal and Torres Strait Islander people because of the cost or time implications of ethical review.

Future monitoring of progress in implementing the Strategy should also examine how ethics review processes are working in practice for evaluations under the Strategy.

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 evaluation reports to be published, subject to privacy or cultural considerations (box 7.3).

Transparency also 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)

Transparency should ... include a commitment to timely publication of reports, including methodological and design aspects, formative and summative, interim and final reports. (The Interplay Project, sub. DR132, p. 5)

That said, transparency is not without some risks (although these can be addressed with quality assurance processes such as external oversight). As the Australian Institute of Primary Care and Ageing said:

There is an important balance between encouraging evaluation and honest uptake and participation and scrutiny over the findings. If there is a requirement that all reports are publicly available, it may discourage frank reporting and create two levels of reporting. This would end up making public information superficial and reduce the contribution it could make to developing public policy. (sub. DR139, 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)

Finally, transparency should be a key principle guiding the governance, monitoring and review of the Strategy itself (chapters 7 and 10).

A Conduct of the project

The Commission actively sought public participation in the development of the Indigenous Evaluation Strategy. This appendix outlines the engagement process and lists the organisations and individuals that 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 oral (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).
- prior to the release of the draft Strategy, the Commission held two roundtables:
 - on 26 November 2019, on the topic of ‘An Indigenous Evaluation Strategy: Objectives, principles and defining success’
 - 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.

In June 2020, the Commission released a draft Strategy to test our thinking and invite further feedback. After the draft Strategy was released, the Commission:

- received a further 68 submissions. These submissions are available online at <https://www.pc.gov.au/inquiries/current/indigenous-evaluation/submissions#post-draft>. One brief comment was also received after the release of the draft Strategy.

-
- undertook further consultations, including with Aboriginal and Torres Strait Islander community controlled organisations and peak bodies, Australian Government agencies, evaluators and academics.
 - held five virtual seminars — three (held on 17 July, 21 July and 23 July 2020) were predominately attended by Aboriginal and Torres Strait Islander people, one held on 22 July 2020 was attended by people outside Australia, and one on 21 August 2020 was attended by people from Australian Government agencies. At each of the seminars the Commission presented an overview of the Strategy and sought feedback from attendees. The Commission also presented at a seminar hosted by the Australian Evaluation Society on August 18 2020.
 - held a virtual priority setting workshop on 14 August 2020 to inform the Commission’s approach to setting evaluation priorities. This was predominately attended by Aboriginal and Torres Strait Islander people.

The Commission thanks all those who contributed to this project.

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 Northern Territory (AMSANT)	81, DR172
Aboriginal Peak Organisations Northern Territory (APO NT)	DR167
Aboriginal Women's Research Assistant and Evaluation Training Project (AWARE)	DR169
AbSec — NSW Child, Family and Community Peak Aboriginal Corporation	09
Accountable Income Management Network (AIMN)	15
AIME Mentoring	37
Alliance of First Nations Independent Education and Training Providers	DR141
Allison, Fiona and LawRight	18
Altman, Jon	23
Anderson, Professor Clair	105 #*
Angelo, Denise; Hudson, Dr Catherine; Macqueen, Dr Susy	76
Animal Management in Rural and Remote Indigenous Communities (AMRRIC)	DR142
ANTaR	DR124
ARC Centre for Excellence for the Dynamics of Language	93
Audiology Australia (AudA)	27, DR147 #
Australian Bureau of Statistics (ABS)	DR133
Australian Council of TESOL Associations (ACTA)	87, DR179
Australian Evaluation Society (AES)	49, DR174
Australian Government Department of Health	35 #
Australian Government Department of the Environment and Energy (DoEE)	101
Australian Indigenous Doctors' Association (AIDA)	DR125
Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)	72
Australian Institute of Health and Welfare (AIHW)	99, DR123
Australian Institute for Primary Care and Ageing (AIPCA)	DR139
Australian Medical Association (AMA)	DR116
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
BetterEvaluation	DR144
Beyond Blue	DR121

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
Bowie, Ian	DR113
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 Australia Academic Health Science Network (CA AHSN)	85 #
Centre for Evidence and Implementation (CEI)	DR143
Children's Ground	73 #
City of Ballarat	54
Clear Horizon	78
Close the Gap Campaign Steering Committee	71, DR159
Commissioner for Children and Young People (WA)	08
Community First Development	DR134
Congress of Aboriginal and Torres Strait Islander Nurses and Midwives (CATSINaM)	01
Couzos, Associate Professor Sophia	92
Cox, Adjunct Professor Eva	68
Cross, Sherrie	82
Curran, Dr Liz	02
CSIRO	DR131
Danila Dilba Health Service	DR165
Darling Downs and West Moreton Primary Health Network	06
Denny-Smith, George; Williams, Megan; Loosemore, Martin	DR153
Department of Health	DR176
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, DR171
Ernst & Young	DR140
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 #
Hassall, Keryn	DR151
Holmwood, Tony	DR175 #

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
Indigenous Allied Health Australia (IAHA)	31
Indigenous Business Australia (IBA)	36, DR120
Indigenous Community Volunteers (ICV)	88
Indigenous Data Network — Indigenous Studies Unit, Uni of Melbourne	32, DR162
Independent Members of the National Indigenous Australians Agency	DR115
Indigenous Evaluation Committee (NIAA IEC) — Distinguished Professor Maggie Walter, Dr Fiona Cram, Dr Anthony Dillion, Dr Wendy Jarvie	
Institute for Human Security and Social Change (IHSSC)	DR119
Institute for Urban Indigenous Health (IUIH)	53
Jobs Australia	57
Just Reinvest NSW	80
Kimberley Aboriginal Medical Services (KAMS)	DR164
Kimberley Aboriginal Law & Cultural Centre (KALACC)	103
Kimberley Aboriginal Law & Cultural Centre (KALACC) and Cultural Development Network (CDN)	DR129
Kimberley Community Legal Services (KCLS)	67
Lateral Economics	DR177
Literacy for Life Foundation	28
Local Government Association of Queensland (LGAQ)	102
Local Government Association of the Northern Territory (LGANT)	39
Lowitja Institute	50, DR157
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, DR148 #
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 (CBPATSISP)	63
National Aboriginal and Torres Strait Islander Legal Services (NATSILS)	97, DR160 #*
National Aboriginal Community Controlled Health Organisation (NACCHO)	95, DR170
National Family Violence Prevention Legal Services Forum (FVPLS)	66
National Centre of Indigenous Excellence (NCIE)	DR158
National Health and Medical Research Council (NHMRC)	89 #
National Health Leadership Forum (NHLF)	59, DR146
National Indigenous Australians Agency (NIAA)	86 #, DR161 #
National Justice Project	51
National Mental Health Commission	19
National Social Security Rights Network (NSSRN)	10
National Voice for our Children (SNAICC)	DR163

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
Ninti One Limited	DR118
NSW Aboriginal Housing Office (AHO)	DR130
NSW Aboriginal Land Council (NSWALC)	91 #, DR122
NSW Ombudsman	DR152 #
NSW Treasury	DR180
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, DR173
Queensland Family and Child Commission	DR136
Queensland Indigenous Family Violence Legal Service (QIFVLS)	25
Queensland Nurses and Midwives' Union (QNMU)	12, DR128
Ragg and Co	DR145
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 #
Riverina Murray Regional Alliance (RMRA)	DR149 #
Royal Australasian College of Physicians (RACP)	77
Royal Australian and New Zealand College of Psychiatrists (RANZCP)	DR138
Saunders, Stuart	DR178
Save the Children	52
Smith, Professor James	47 #
Social Ventures Australia (SVA)	83, DR168
Sullivan, Dr Patrick; Hunt, Dr Janet; Lahn, Dr Julie	11 #
Tasmanian Government	100, DR166
Tauondi Aboriginal Corporation (trading as Tauondi Aboriginal College)	94
The Fred Hollows Foundation (TFHF)	14, DR156
The Interplay Project Pty Ltd	DR132 #
The Law Society NSW	DR155
The Smith Family	56, DR154
The University of Queensland	20
Torres Shire Council	108

(continued next page)

Table A.1 (continued)

<i>Participants</i>	<i>Submission number</i>
Tranby National Indigenous Adult Education and Training	96
UNSW Social Policy Research Centre	DR127
Vic Department of Health and Human Services (DHHS)	DR137
Victorian Aboriginal Child Care Agency (VACCA)	26, DR126
Victorian Aboriginal Community Controlled Health Organisation (VACCHO)	44
Vision 2020 Australia	DR135
Walgett Aboriginal Medical Service (WAMS)	04
Walter, Professor Maggie — Oral submission	112
Weatherburn, Professor Don	DR114
Western Australian Council of Social Services (WACOSS)	21
Wuchopperen Health Service Limited	DR117
Western Australian Government	74 #
Williams, Dr Megan and Ragg, Dr Mark	45
Winnunga Nimmityjah AHCS	46
Yothu Yindi Foundation (YYF)	60 #
Yulang Indigenous Evaluation	DR150
Yuwaya Ngarra-li partnership, Dharrwaa 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 (AIATSIS)
Australian Institute of Health and Welfare
Australian National Audit Office
Australian Public Service Commission
Australian Government Treasury (The Treasury)
Bissell, Michael
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
University of Canberra
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-binya 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

(continued next page)

Table A.2 (continued)

Participants

Consultations held by videoconference (after the release of the draft Strategy)

Australian Evaluation Society
Australian Government Treasury (The Treasury)
Australian Indigenous Doctors' Association (AIDA)
Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)
Australian Institute of Health and Welfare
Australian Public Service Commission
Bissell, Michael
Cargo, Professor Margaret
Clear Horizon Consulting
Coalition of Aboriginal and Torres Strait Islander Peak Organisations
Community First Development
Cultural and Indigenous Research Centre Australia (CIRCA)
Department of Education, Skills and Employment
Department of Finance
Department of Foreign Affairs and Trade
Department of Health
Department of Industry, Science, Energy and Resources
Department of Social Services
Department of the Prime Minister and Cabinet
Empowered Communities Leadership Group
Ernst & Young Evaluation Practice Network / Indigenous Sector Practice
Finlay, Dr Summer
Fredericks, Professor Bronwyn
Gruen, Nicholas, Lateral Economics
Karabena Consulting
Indigenous Allied Health Australia
Inside Policy
IPSOS Aboriginal and Torres Strait Islander Research Unit
National Centre of Indigenous Excellence (NCIE)
National Indigenous Australians' Agency (NIAA)
Ninti One Limited
Nous Group
PwC Indigenous Consulting
Remote Medicine Academy
Ridgeway, Dr Aden
Winangali Limited

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 Leadership Group
Indigenous Allied Health Australia
Lowitja Institute
National Aboriginal and Torres Strait Islander Health Worker Association
National Aboriginal and Torres Strait Islander Legal Services
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 – NSW 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

(continued next page)

Table A.3 (continued)

14 August 2020 – ‘Priority-setting Workshop’

Aborigines Advancement League
Australian Healthcare and Hospitals Association
Australian Human Rights Commission
Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)
Coalition of Aboriginal and Torres Strait Islander Peak Organisations
Department of the Prime Minister and Cabinet
Empowered Communities Leadership Group
Fredericks, Professor Bronwyn
Gayaa Dhuwi (Proud Spirit) Australia
Lowitja Institute
National Aboriginal and Torres Strait Islander Health Worker Association
National Aboriginal and Torres Strait Islander Higher Education Consortium
National Aboriginal and Torres Strait Islander Legal Services
National Family Violence Prevention and Legal Services Forum
National Native Title Council
National Indigenous Australians Agency (NIAA)
NSW Aboriginal Land Council (NSWALC)
SNAICC – National Voice for Our Children
Torres Strait Regional Authority
Victorian Aboriginal Child Care Agency (VACCA)
Walter, Professor Maggie
Yu, Professor Peter

B Information request to Australian Government agencies

B.1 About the information request

To inform our understanding of current evaluation policies and practices in Australian Government agencies, the Commission sent an information request to 182 agencies in November and December 2019. Agencies were identified using the *List of Commonwealth entities and companies under the Public Governance, Performance and Accountability Act 2013*.⁶⁶

The information request asked questions about evaluation generally and evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people, including:

- mainstream policies and programs — policies and programs designed for all eligible Australians, such as most social security payments and the Medicare Benefits Schedule
- Indigenous-specific policies and programs — policies and programs designed specifically for Aboriginal and Torres Strait Islander people, such as the National Indigenous Australians Health Program and the Indigenous Advancement Strategy.

The information request asked agencies for:

- details on their evaluation and data policies, evaluation planning and governance arrangements and evaluation practices
- a list of evaluations undertaken in the years 2016-17 to 2019-20.⁶⁷

Changes to Australian Government departments came into effect in February 2020. We asked that responses reflected the situation at December 2019 prior to changes taking effect.

This appendix:

- documents response rates, methods used to analyse data, and limitations (section B.2)
- summarises agencies' responses to the questions in the information request (section B.3)

⁶⁶ Information requests were initially also sent to the three parliamentary departments — Department of Parliamentary Services, Department of the House of Representatives and Department of the Senate — and the Parliamentary Budget Office. However, these agencies were subsequently removed from the sample as they are departments and agencies of the Parliament rather than the Government, and outside the scope of the Indigenous Evaluation Strategy.

⁶⁷ Further analysis of published and unpublished evaluation reports provided by agencies was undertaken to collect data on evaluation characteristics.

- analyses data collected by the Commission from evaluation reports that agencies provided in their responses to the information request (section B.4)
- reproduces the questionnaire sent to agencies (section B.5).

B.2 Methodology and data quality

Responses received

52 agencies (or 29 per cent of agencies) responded to the information request (table B.1).⁶⁸

Table B.1 Australian Government agencies that responded to the information request

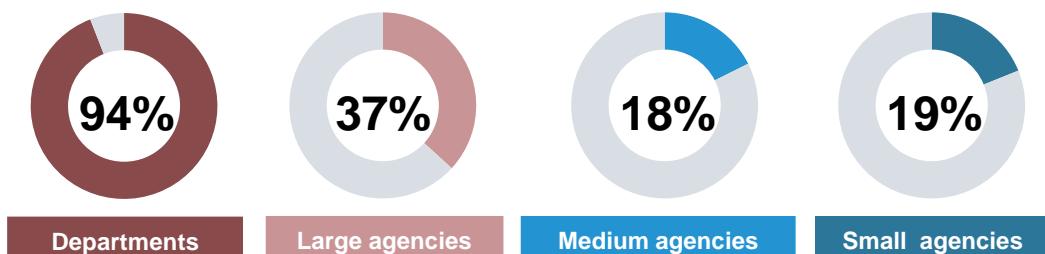
Departments	
Attorney General's Department	Department of Foreign Affairs and Trade
Department of Agriculture	Department of Health
Department of Communications and the Arts	Department of Home Affairs
Department of Defence	Department of Industry, Innovation and Science
Department of Education	Department of Infrastructure, Transport, Cities and Regional Development
Department of Employment, Skills, Small and Family Business	Department of Social Services
Department of the Environment and Energy	Department of the Prime Minister and Cabinet
Department of Finance	Department of the Treasury
Other agencies	
Aboriginal Hostels Ltd	Comcare
AgriFutures Australia	Food Standards Australia New Zealand
Army Relief Trust Fund	Future Fund Management Agency
Australia Council for the Arts	Great Barrier Reef Marine Park Authority
Australian Curriculum, Assessment and Reporting Authority	Indigenous Business Australia
Australian Electoral Commission	Infrastructure and Project Financing Agency
Australian Federal Police	Inspector-General of Taxation and Taxation Ombudsman
Australian Naval Infrastructure Pty Ltd	IP Australia
Australian Pesticides and Veterinary Medicines Authority	Murray-Darling Basin Authority
Australian Renewable Energy Agency	National Archives of Australia
Australian Research Council	National Indigenous Australians Agency
Australian Sports Commission (Sports Australia)	National Transport Commission
Australian Taxation Office	NDIS Quality and Safeguards Commission
Australian Trade and Investment Commission	Office of the Australian Information Commissioner
Australian Transport Safety Bureau	Organ and Tissue Authority
Bureau of Meteorology	Outback Stores
Coal Mining Industry (Long Service Leave Funding) Corporation	Safe Work Australia
	Screen Australia
	Services Australia

Source: Productivity Commission Indigenous Evaluation Strategy information request.

⁶⁸ Agency names in table B.1 are those that were in use as at December 2019.

The Commission received responses from all but one department (figure B.1). The response rate from other agencies was 22 per cent. Response rates were lower for small- and medium-sized agencies than for large agencies (figure B.1).

Figure B.1 Agencies' response rate for the information request, by agency type and size^a



^a Agencies (other than departments) are categorised into three groups by 2018-19 estimated total resourcing: small (less than \$30 million); medium (\$30-\$200 million); large (more than \$200 million). Excludes agencies for which data on resourcing were not available; response rate for these agencies was 13 per cent.

Data sources: Productivity Commission Indigenous Evaluation Strategy information request; Department of Finance, 2019-20 *Parliamentary Budget Statement Agency tables*, accessed at data.gov.au.

The sample of departments is large enough to be confident that the data collected accurately reflects evaluation policies and practices in departments. However, the lower response rate for other agencies does not preclude the possibility of response bias. Those agencies that are well organised from an evaluation perspective, or felt they had a good story to tell, could have been more likely to respond to our information request. This limitation should be kept in mind when interpreting the results.

Analysis of questionnaire responses

Data from agencies' responses to questions about evaluation policies and practices were compiled and presented in descriptive charts and tables (section B.3). Where possible, and where results differ across types of agencies, results are presented separately for departments and other agencies.

Some agencies failed to answer or answered 'not applicable' to some of the questions. This was typically when an agency that reported doing no evaluation was asked about evaluation practices. For example, there was a relatively large number of missing responses to questions about planning for evaluation when new policies and programs are developed. This is likely to be due to some responding agencies having no responsibility for developing new policies and programs. Likewise, agencies that reported doing no evaluation in recent years typically did not provide answers to questions about ethics and publication practices. In these cases,

results are presented for the subset of agencies that answered the question. Notes to tables and figures provide more detail.

Some of the questions allowed open-ended responses. In order to compare across agencies, the Commission summarised these data into categories. However, the coding of responses into categories involved a degree of subjectivity. The questions involved concerned:

- decisions about what policies and programs get evaluated, including criteria for deciding on evaluation priorities and how often evaluation priorities are identified
- procedures for sharing evaluation results within agencies.

Data received from question five of the information request (relating to evaluation governance arrangements) were difficult to compare accurately across agencies. Many smaller agencies identified audit or risk committees when asked if they had an agency-wide committee to oversee evaluation. However, several larger agencies and departments said that they had no committee to oversee evaluation, despite being likely to have similar audit and risk committees. The intent of this question was to identify governance mechanisms that were specifically designed to oversee evaluation within agencies, rather than general agency governance arrangements. Responses from these questions were not used in the analysis in section B.3.

Analysis of evaluation reports

The information request asked agencies to provide copies of, or links to, evaluations that they had conducted or commissioned from 2016-17 to 2019-20. Where evaluation reports were unpublished, agencies were asked to provide copies in confidence. Where this was not possible, agencies were asked to estimate the number of unpublished evaluation reports for each year.

Data from the information request were used to estimate the scale of evaluation activity in Australian Government agencies. Evaluation reports provided to the Commission were also analysed to gain an understanding of evaluation practices, including:

- the extent to which evaluations assessed impacts on Aboriginal and Torres Strait Islander people
- the use of external consultants
- methods and data used
- engagement with Aboriginal and Torres Strait Islander people during evaluation
- ethical review
- use of evaluation findings (section B.4).

Much of the analysis focuses on a subset of evaluation reports that mention or provide results for Aboriginal and Torres Strait Islander people. These are divided into two categories:

- mainstream evaluations with Indigenous results — these are evaluations of mainstream policies and programs that mention or provide results for Aboriginal and Torres Strait Islander people
- Indigenous-specific evaluations — these are evaluations of Indigenous-specific policies and programs.

Conclusions drawn from analysis of evaluation reports should be used with some caution. In addition to concerns raised above about the representativeness of the sample, issues include:

- *not all evaluations were able to be identified* — some agencies acknowledged that the evaluation reports they provided were a subset of evaluations they had undertaken because they were unable to identify all evaluations to include in their response (this included agencies where responsibility for evaluation was devolved to program areas, and/or where there was not a central register of evaluations)
- *evaluations may not be comparable* — the nature and size of evaluations, and the policies and programs being evaluated, varies. This means that counting the numbers of evaluations with different characteristics is only a rough approximation of patterns of evaluation practice
- *evaluation conduct may not be reflected in evaluation reports* — assessment of evaluation conduct is based on information available in evaluation reports, such as details on methodology, data, and engagement. Reports reflect the perspective of the agency commissioning the evaluation or the evaluator, rather than necessarily the perspectives of users, community members, service providers or experts on the usefulness or quality of the evaluation.

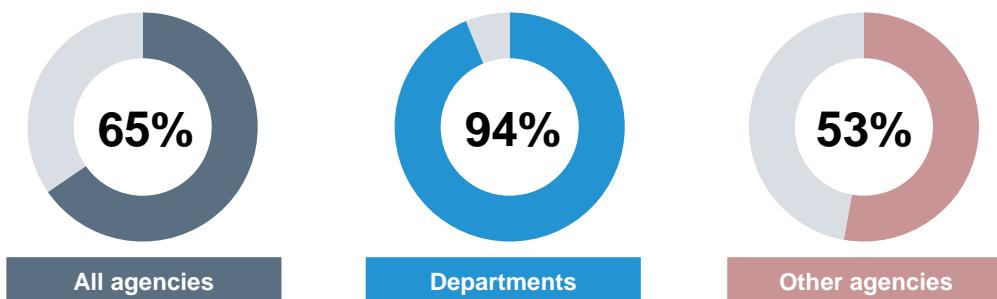
Analysis of other documents

Agencies responding to the information request were asked to provide copies of various policy documents. The Commission used these documents to analyse agencies' evaluation policies, strategies and plans (chapter 2) and data strategies and plans (chapter 9).

B.3 Evaluation policies and practices

Most departments and more than half of the other agencies reported having planned, conducted or commissioned evaluations between 2016-17 and 2019-20 (figure B.2). Further details on the number and characteristics of evaluations undertaken by agencies are in section B.4.

Figure B.2 Agencies that planned, conducted or commissioned evaluations in the years 2016-17 to 2019-20^a



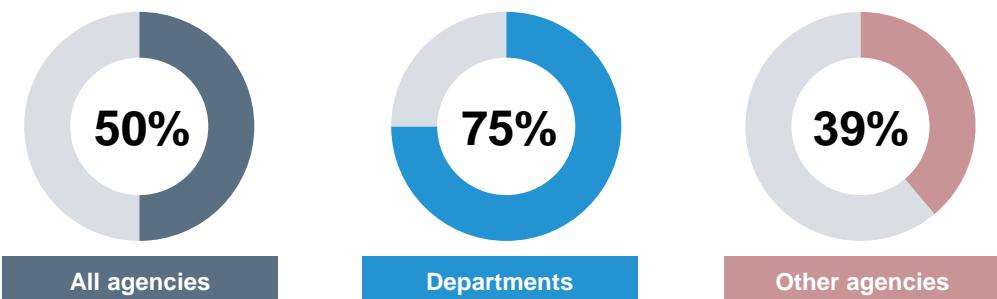
^a Based on responses from 52 agencies: 16 departments and 36 other agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 2).

Evaluation and data policies

Half of all agencies reported that they had a formal evaluation policy, framework, strategy, plan or guide (figure B.3). Departments were more likely to have evaluation policies, strategies or guidance on evaluation activities than other agencies.

Figure B.3 Agencies with formal policies, frameworks, strategies, plans or guides for evaluation activity^a



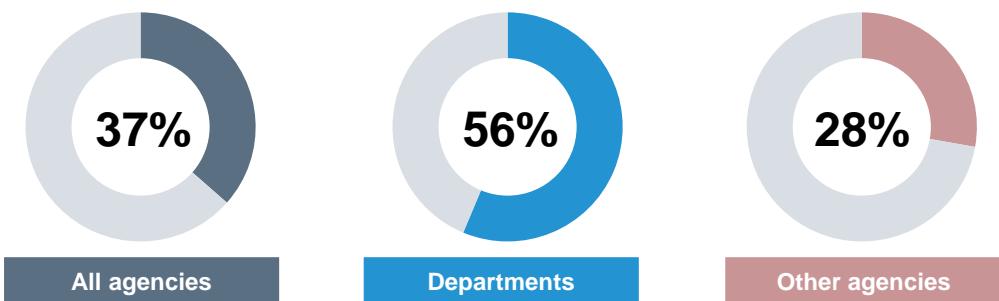
^a Based on responses from 52 agencies: 16 departments and 36 other agencies.

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

Departments were more likely than other agencies to report having a formal data policy or strategy (figure B.4). To ensure consistency across responses, agencies that reported having a privacy policy but no other formal data policy or strategy were not included in the total.⁶⁹

⁶⁹ Several smaller agencies reported that their privacy policy was their data policy or strategy. Few departments reported privacy policies in this item, however all Australian Government agencies with turnover greater than \$3 million are required to have a privacy management plan under the *Privacy*

Figure B.4 Agencies with a formal policy, strategy or guidelines for collecting and managing data^a



^a Based on responses from 52 agencies: 16 departments and 36 other agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 13).

The information request asked whether agencies had a formal data policy, strategy or guidelines for collecting and managing data about Aboriginal and Torres Strait Islander people, and to provide a copy of the relevant policy to the Commission. Responses to this question varied in quality: many agencies identified privacy policies or human resources policies relating to recruitment of Aboriginal and Torres Strait Islander staff.

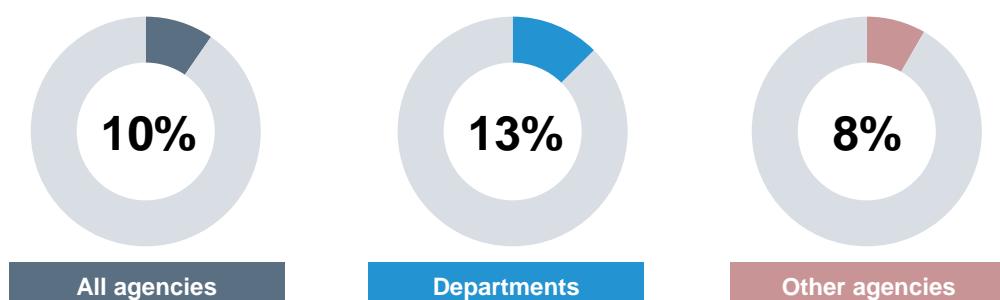
Rather than use the questionnaire responses, the Commission examined the policy documents provided to determine whether data policies were specifically for data about Aboriginal and Torres Strait Islander people, or whether general data policy documents mentioned Aboriginal and Torres Strait Islander data.⁷⁰ Agencies that reported having neither a general data policy of any kind nor an Indigenous-specific data policy were classified as not having a data policy that considers Aboriginal and Torres Strait Islander data.

The results show that very few agencies of any size have a formal data policy or strategy that considers Aboriginal and Torres Strait Islander data (figure B.5).

(Australian Government Agencies – Governance) APP Code 2017. As such, it is likely that most agencies responding to the information request have a privacy policy, even if it was not mentioned in their response.

⁷⁰ General data policies were searched for the terms ‘Aboriginal’, ‘Torres Strait Islander’ and ‘Indigenous’.

Figure B.5 Agencies with a formal data policy or strategy that considers Aboriginal and Torres Strait Islander data^a

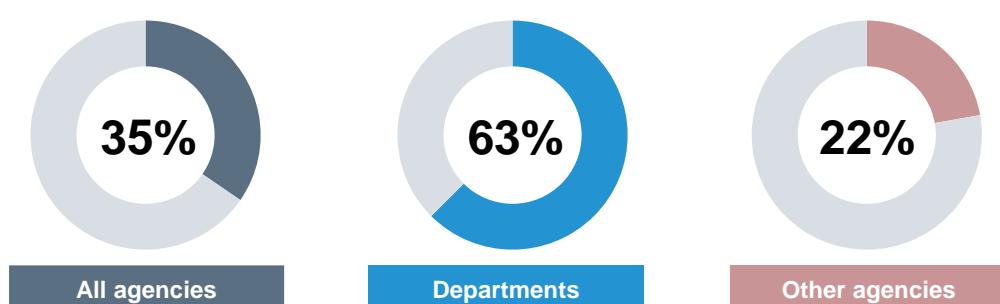


^a Based on analysis of data policy documents provided by 52 agencies: 16 departments and 36 other agencies.
Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 14).

Central evaluation units

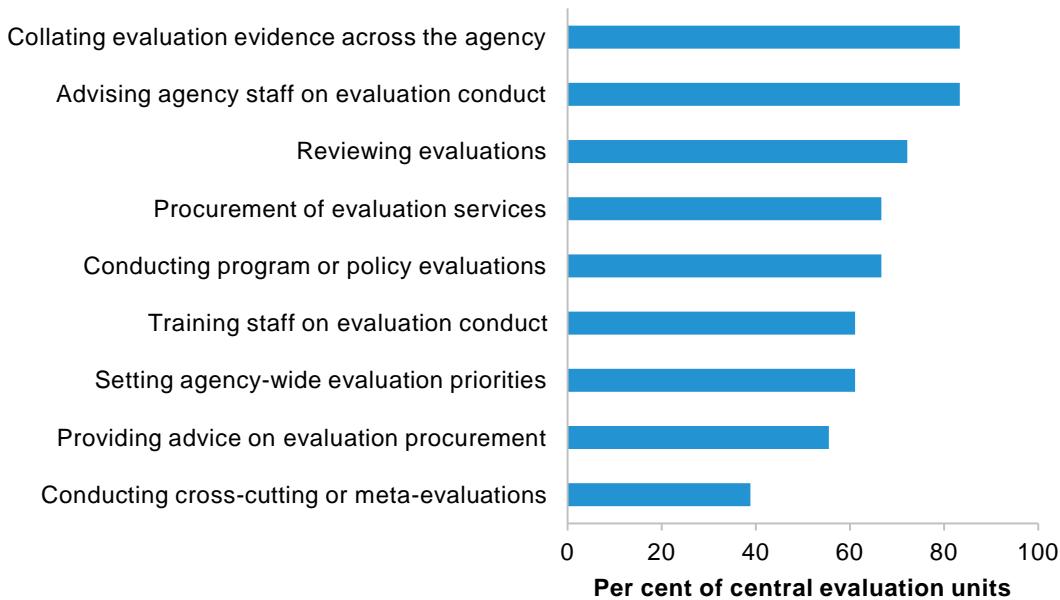
More than one-third of agencies reported having a unit dedicated to evaluation of policies and programs (figure B.6), with units more likely to be found in departments than other agencies. Among the 18 central evaluation units within responding agencies, most collated evaluation evidence and advised staff on evaluation conduct. Almost three-quarters reviewed evaluations conducted or commissioned by the agency. Two-thirds of central evaluation units were involved in procuring evaluation services and/or undertaking evaluation themselves. Only 39 per cent conducted cross-cutting or meta-evaluations (figure B.7).

Figure B.6 Agencies with a central evaluation unit^a



^a Based on responses from 52 agencies: 16 departments and 36 other agencies.
Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 4).

Figure B.7 Roles of central evaluation units^a



^a Based on responses from 18 agencies that reported having a central evaluation unit. Agencies were asked to select roles from a list of possible roles. More than one response could be selected by each agency.

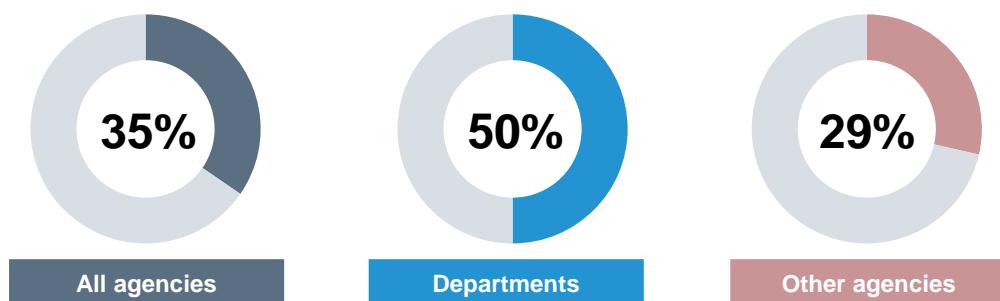
Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 4).

Evaluation planning

About 35 per cent of agencies reported having a set of formal criteria used for determining which policies and programs are evaluated (figure B.8). Agencies were only included in this total if their response listed a set of criteria or pointed to evaluation policy documents that listed a set of criteria used in a formal process of determining evaluation priorities.

Of those agencies that listed a set of criteria for deciding which policies and programs are evaluated, the most commonly used criteria were policy or program size or expenditure, and strategic significance (figure B.9). The impact and risk associated with a policy or program were also considered relatively frequently. A small number of agencies reported that all policies and programs were evaluated on a rolling basis, or that decisions of government or legislative requirements determined which policies and programs were evaluated. Other less commonly-cited criteria included where a policy or program has a high profile, where there is a lack of evidence (particularly in the case of pilot programs), or where there is a planned change in the function or funding for a policy or program. In some cases, evaluations are triggered by evidence of poor outcomes from monitoring or performance data.

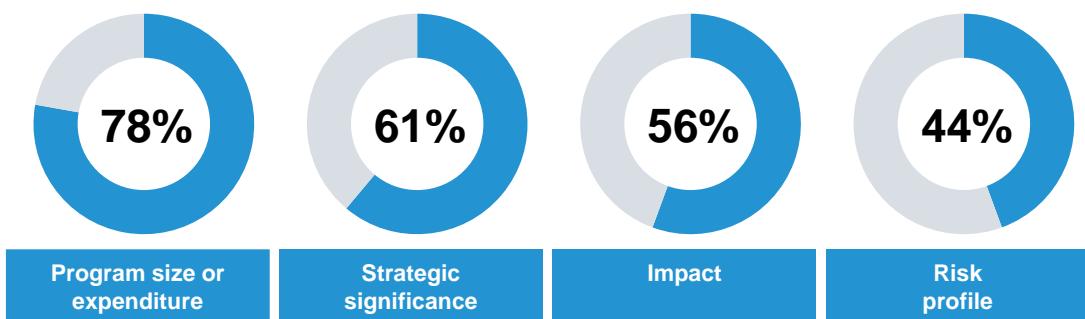
Figure B.8 Agency has formal criteria for deciding which policies and programs are evaluated^a



^a Based on responses from 52 agencies: 16 departments and 36 other agencies. Open-ended responses have been coded into categories by the Commission.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 6).

Figure B.9 Criteria used for deciding which policies and programs are evaluated, per cent of agencies^a

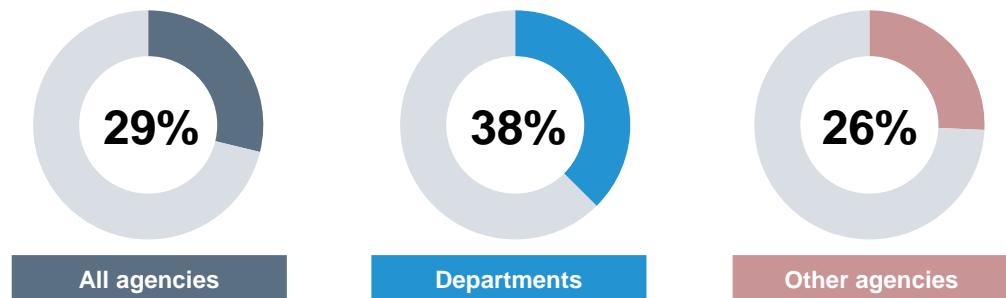


^a Based on responses from 18 agencies that reported having formal criteria for deciding which policies and programs are evaluated. Open-ended responses have been coded into categories by the Commission.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 6).

About 29 per cent of agencies have processes they use annually to identify policies and programs for evaluation (figure B.10). However, in some agencies, this is part of a general business planning or audit exercise rather than being specifically to identify and prioritise policies and programs for evaluation. In most agencies, there is either no process of identifying evaluation priorities, or it is done on an ad hoc basis.

Figure B.10 Agencies that identify evaluation priorities annually^a

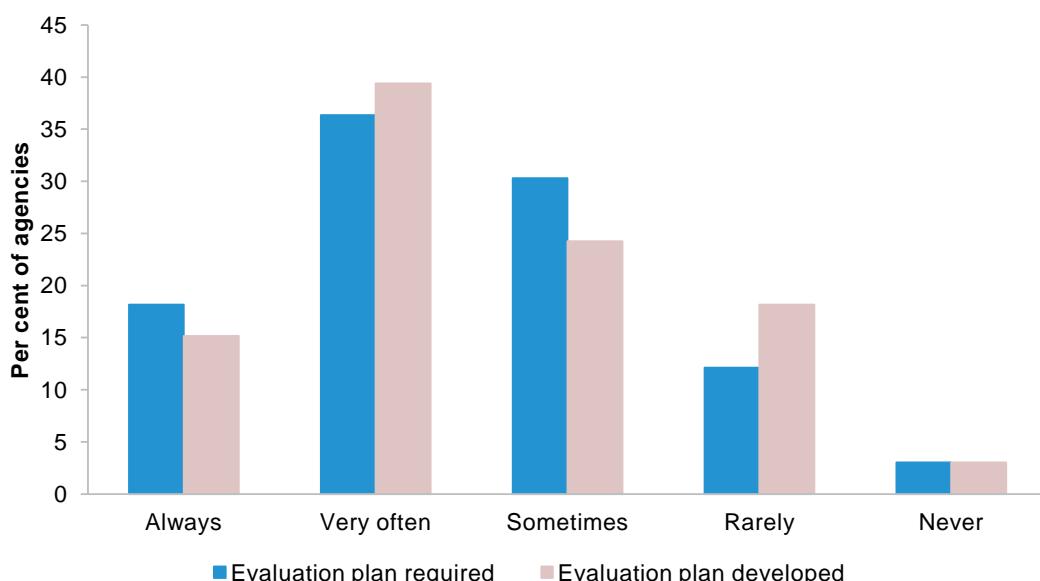


^a Based on responses from 52 agencies: 16 departments and 36 other agencies. Open-ended responses have been coded into categories by the Commission.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 8).

Over half of the agencies reported that evaluation plans were always, or very often, required when developing new policies and programs (figure B.11). About 21 per cent of agencies reported that evaluation plans were never or rarely developed when developing new policies and programs. Early planning for evaluation is more likely to occur in departments than other agencies (figure B.12).

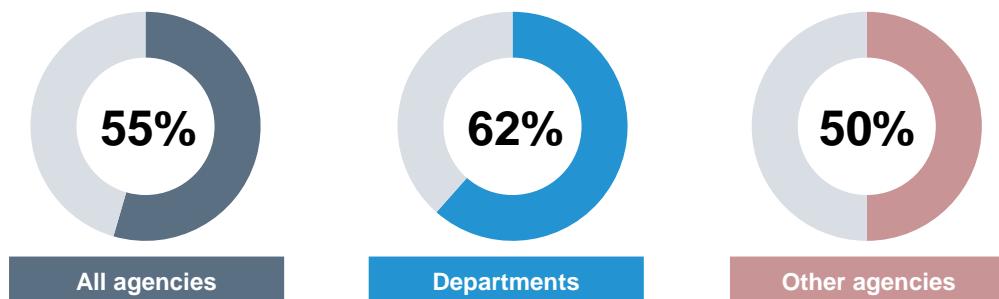
Figure B.11 How often are evaluation plans required and developed for new policies and programs?^a



^a Based on responses from 33 agencies. Excludes agencies that answered 'not applicable', typically those that do not develop new policies and programs, and agencies that did not answer both parts of this question.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 9).

Figure B.12 Agencies that report ‘always’ or ‘very often’ developing evaluation plans when new policies and programs are developed^a



^a Based on responses from 33 agencies: 13 departments and 20 other agencies. Excludes agencies that answered ‘not applicable’ to questions about evaluation planning, typically those that do not have responsibility for developing new policies and programs, as well as agencies that did not answer both parts of this question.

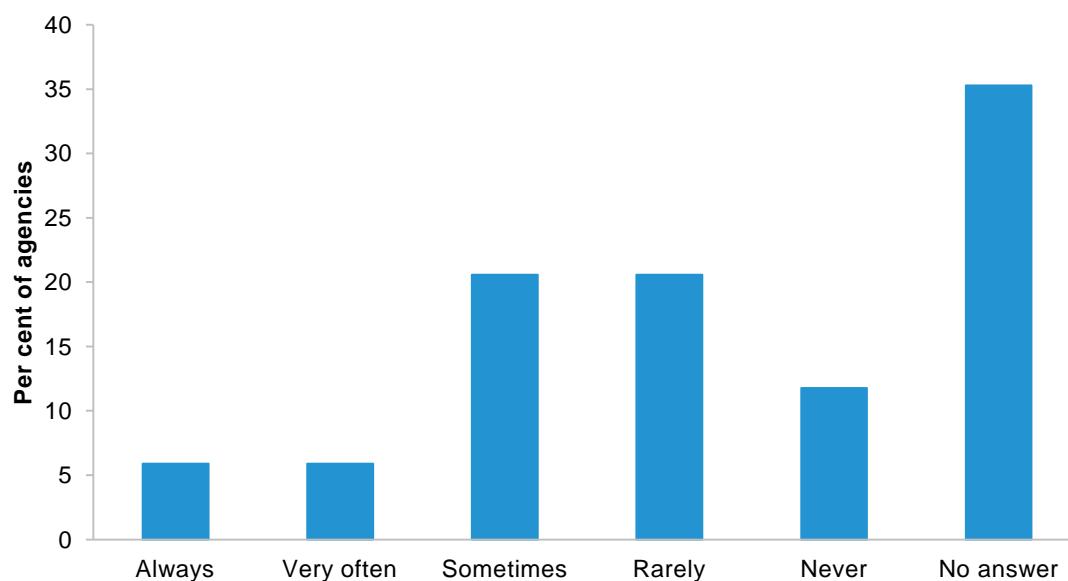
Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 9).

Ethical review

Few agencies reported regularly seeking formal ethical review during evaluations (figure B.13). Only 12 per cent of agencies that reported doing evaluation in the years 2016-17 to 2019-20 said that they always or often sought ethical review for evaluations, and less than a third reported always, very often or sometimes seeking ethical review. This is likely to be due in part to the nature of evaluations undertaken — agencies that reported seeking ethical review more frequently were typically social policy agencies where evaluations are more likely to involve vulnerable participants. Departments were more likely than other agencies to always, very often or sometimes seek ethical review (figure B.14).

About 35 per cent of agencies that reported doing evaluation since 2016-17 did not provide an answer for this question. This may be because they considered that the question was not relevant due to the nature of evaluations done. Regardless, it is clear that ethical review is not frequently considered by agencies.

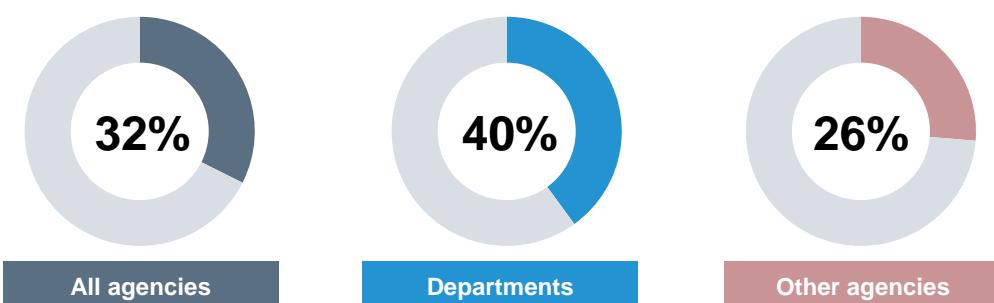
Figure B.13 How often is formal ethical review sought for evaluation activities conducted or commissioned by your agency?^a



^a Based on responses from 34 agencies that reported conducting or commissioning evaluations in the years 2016-17 to 2019-20.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 10).

Figure B.14 Agencies that report 'always', 'very often' or 'sometimes' seeking formal ethical review for evaluation activities they conduct or commission^a



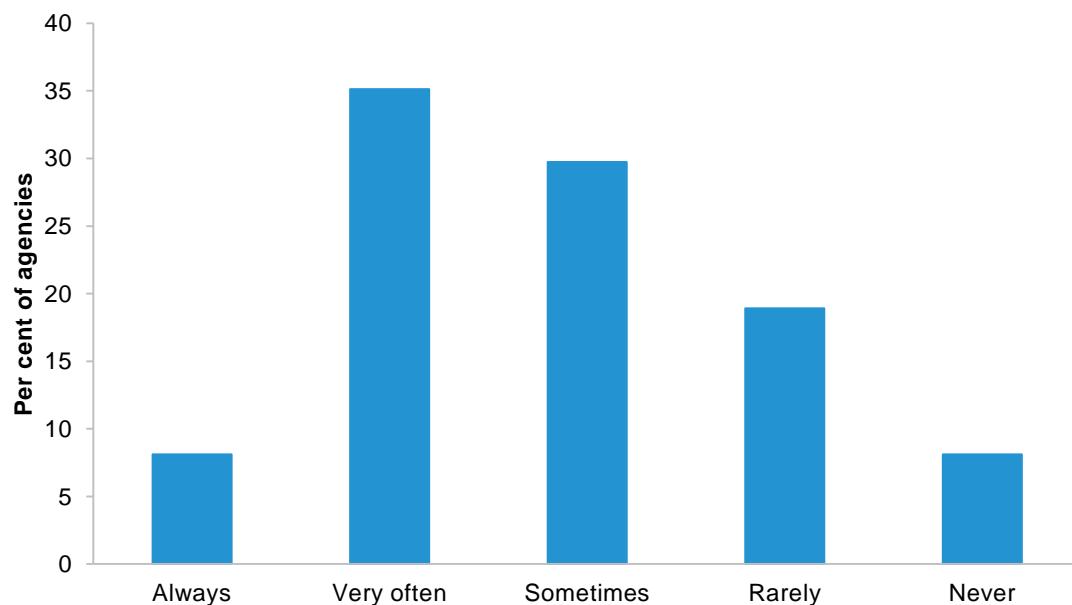
^a Based on responses from 34 agencies that reported conducting or commissioning evaluations in the years 2016-17 to 2019-20: 15 departments and 19 other agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (question 10).

Publication of evaluation reports

Agencies were asked whether evaluation reports are typically published on the agency's external website. Publication practices varied across agencies (figure B.15). Only 8 per cent of agencies said that they always published evaluation reports, with 35 per cent saying that evaluation reports were very often published. Departments are more likely than other agencies to regularly publish evaluation reports, although none of the departments responding to the questionnaire said that they always publish evaluation reports (figure B.16).

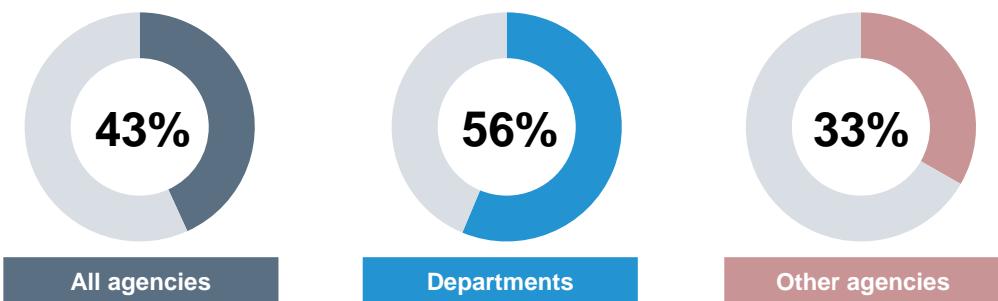
Figure B.15 How often are evaluation reports made available on agencies' external websites?^a



^a Based on responses from 37 agencies. Sample excludes agencies that answered 'not applicable' to this question, which are typically agencies that report doing no recent evaluation.

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

Figure B.16 Agencies reporting that evaluation reports are ‘always’ or ‘very often’ published on the agency’s external website^a



^a Based on responses from 37 agencies: 16 departments and 21 other agencies. Sample excludes agencies that answered ‘not applicable’ to this question, which are typically agencies that report doing no recent evaluation.

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

B.4 Characteristics of evaluations

Most evaluations identified by agencies were unpublished

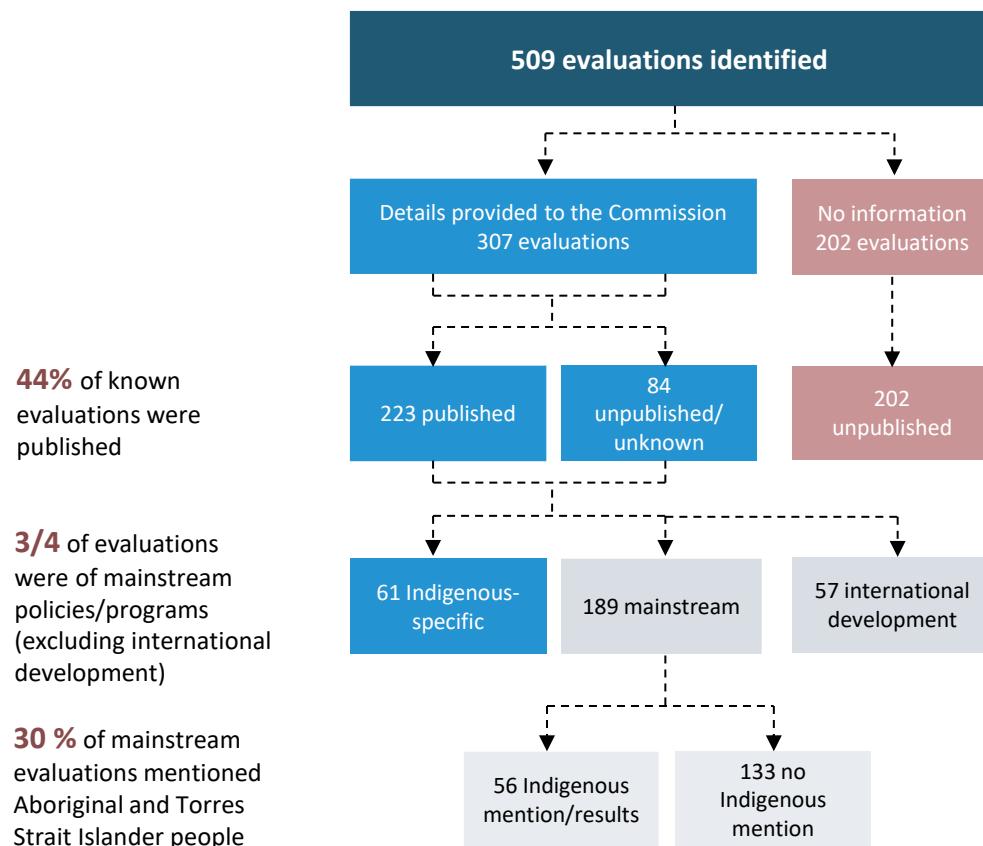
Forty-four per cent of the 509 Australian Government evaluations identified through the Commission’s information request were published.

Agencies provided specific information on 307 evaluations conducted between 2016-17 and 2019-20. Nearly three-quarters of these 307 evaluations were published, and one-quarter were unpublished. Agencies also identified (but did not provide details on) a further 202 unpublished evaluations (figure B.17).

Evaluation numbers vary across agencies and service areas

Numbers of evaluations undertaken varied across agencies. Based on our sample of responses, the number of evaluations undertaken does not seem to be correlated to agency size. Some small agencies conducted a number of evaluations, while others undertook very little or no evaluation. However, 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.

Figure B.17 Known evaluations by Australian Government agencies^a



^a Counts represent information provided by the 52 agencies that responded out of 182 agencies the request was sent to. 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.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

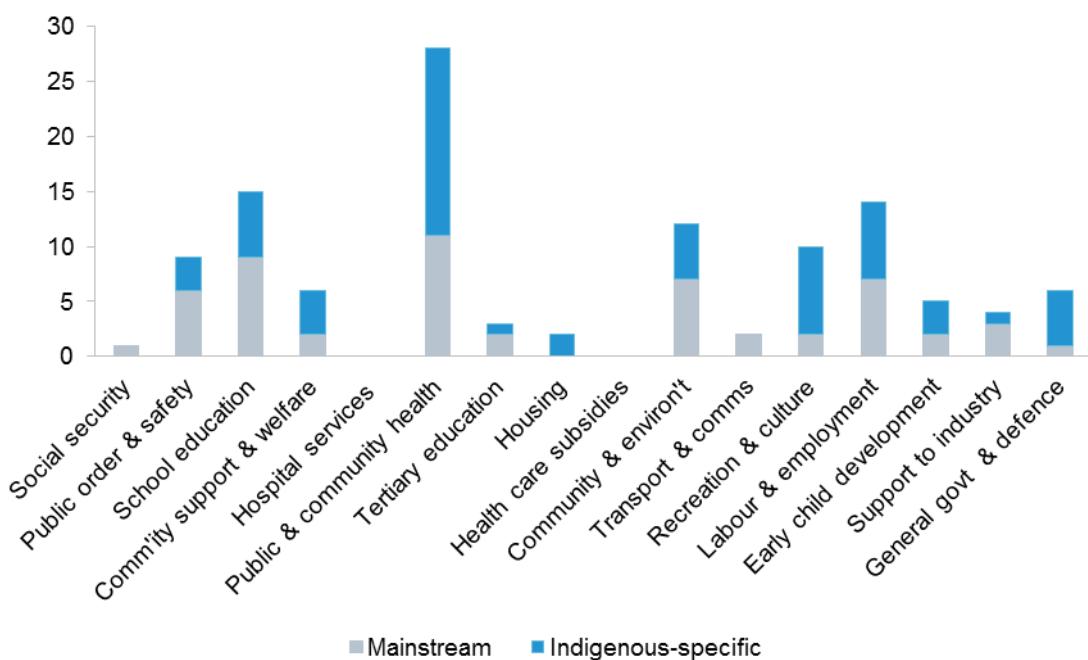
Of the evaluations given to the Commission, 57 were evaluations of international development assistance programs commissioned by the Department of Foreign Affairs and Trade. These evaluations examine programs implemented in other countries and have been excluded from the analysis of evaluations in this section to focus on policies and programs implemented in Australia (250 evaluations). Three-quarters (189 out of 250) of these evaluations were for mainstream policies and programs. However, only 30 per cent (56) of these mentioned or provided results specific to Aboriginal and Torres Strait Islander people (figure B.17). There were 61 evaluations of Indigenous-specific policies or programs.

Most of the data drawn from the information request in the remainder of this section are focussed on evaluations that mention or provide results on Aboriginal and Torres Strait Islander people. Some of these evaluations provide results on outcomes for Aboriginal and Torres Strait Islander people, while others provide more limited information (including, for

example, the disaggregation of some data by Indigenous status or acknowledging possible effects of policies and programs on Aboriginal and Torres Strait Islander people).

The numbers of evaluations of policies and programs mentioning or providing results on Aboriginal and Torres Strait Islander people varied substantially across service areas. A quarter of these were for public and community health services. Other areas where evaluation was more common were school education, labour and employment, and community and environment services (figure B.18). Areas where there were few or no recent evaluations were social security payments and healthcare subsidies and support (including Medicare and the Pharmaceutical Benefits Scheme) — coincidentally these are areas of significant government expenditure (chapter 3).

Figure B.18 Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people, by service area^{a,b}



^a Service areas match the definitions used in the *Indigenous Expenditure Report* (SCRGSP 2017). ^b Data should be interpreted with caution as not all agencies provided data, some were not able to report on all their evaluations and about one third of known evaluations were unpublished with no information on service area available.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

Two-thirds of mainstream evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people related to government programs, the rest related to policies. For Indigenous-specific evaluations, more than 90 per cent were for government programs (table B.2).

Very few evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people examined issues across agencies or service areas — three Indigenous-specific policy evaluations and one mainstream program evaluation covered multiple agencies or service areas.

Evaluations of government policies were not always described as ‘evaluations’ but more commonly referred to as ‘reviews’. These are included in this assessment as evaluations as they share the same evaluative focus of assessing effectiveness, efficiency and appropriateness. Some policy evaluations were overseen by independent external committees, and some sought public submissions as part of their research processes. Productivity Commission inquiries were not included as evaluations in the data presented here, but many of them share characteristics with policy evaluations.

Table B.2 Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people

	Mainstream	Indigenous-specific	Total
Program evaluations ^a	38	56	94
Policy evaluations ^a	18	5	23
Total evaluations	56	61	117

^a Count of evaluations. Program evaluations are evaluations of a program, service, or payment by an agency. Policy evaluations are evaluations of policies that do not involve delivering a program, payment or service. Policies may relate to laws, regulations, taxes, charges or administrative requirements imposed on individuals, firms or government agencies.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

Mixed methods evaluation techniques dominate

About two-thirds of evaluations mentioning or providing 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 (table B.3). About one in seven evaluations were based solely on quantitative methods. The most common evaluation methods used were:

- literature reviews and analysis of documents
- 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.

Table B.3 Evaluation methods used^a

	<i>Mainstream</i>	<i>Indigenous-specific</i>	<i>Total</i>
Mixed methods ^b	39	28	67
Only quantitative methods	8	6	14
Only qualitative methods	9	10	19
Unknown ^c	—	17	17
Total evaluations	56	61	117

^a Count of evaluations. ^b Evaluations using a mix of quantitative and qualitative methods. ^c There was insufficient information about methods in some evaluations to classify them.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

Qualitative techniques are useful for evaluating why a policy or program is working or not and why it may be more or less effective for particular situations or types of service users (chapter 4).

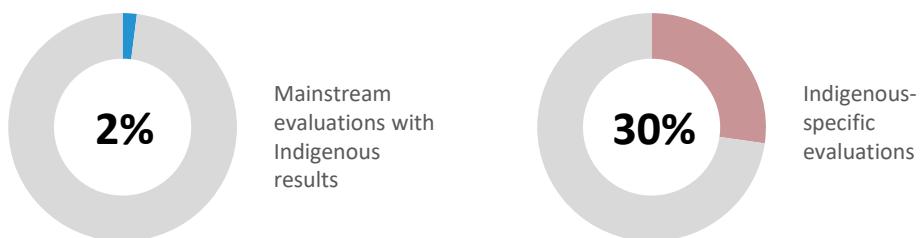
Quantitative analysis in Australian Government evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people primarily involved analysis of administrative data and surveys. Some evaluations had good data on outcomes, however, the majority of quantitative analysis focused on activities, outputs, and user perceptions and feedback. While these can all be useful in their own way, they can be of limited value for measuring outcomes.

Most evaluations attempt to measure impact, but most have limited data and/or do not have a control group. Only four evaluations of policies or programs mentioning or providing results for Aboriginal and Torres Strait Islander people included economic analysis. (Chapter 9 provides more information on the use of data in evaluations.)

Engagement with Aboriginal and Torres Strait Islander people is minimal

About 30 per cent of evaluations of Indigenous-specific policies or programs included engagement with Aboriginal and Torres Strait Islander people in evaluation planning and decision making. Only one out of 56 mainstream evaluations included Aboriginal and Torres Strait Islander people in planning and decision making (figure B.19). The lack of engagement was not restricted to evaluations of policies or programs affecting Aboriginal and Torres Strait Islander people — for evaluations more generally, most or all decisions about evaluation design and governance were made without input from service users, service providers, community organisations or other external parties.

Figure B.19 Evaluations that included Aboriginal and Torres Strait Islander people in planning and decision-making^a



^a Proportion of evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people that included Aboriginal and Torres Strait Islander people in evaluation planning or decision-making.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

For evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people, participation of service users and providers in evaluations was primarily limited to interviews, group discussions or as survey respondents. More than half (30 out of 56) of mainstream evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people had no engagement with Aboriginal and Torres Strait Islander people; about one-third (18) had some engagement, and for the remainder (8) there is insufficient information on the level of engagement. Most of the mainstream evaluations without any direct engagement presented administrative or other data on Aboriginal and Torres Strait Islander people.

About half (32 out of 61) of Indigenous-specific evaluations engaged with Aboriginal and Torres Strait Islander people as part of the research process, mostly in the form of interviews, group discussions and as survey respondents. Seven out of 61 Indigenous-specific evaluation reports mentioned having Aboriginal and Torres Strait Islander people as part of the research or evaluation team. For the remaining evaluations, most have insufficient information to know whether there was engagement with Aboriginal and Torres Strait Islander people.

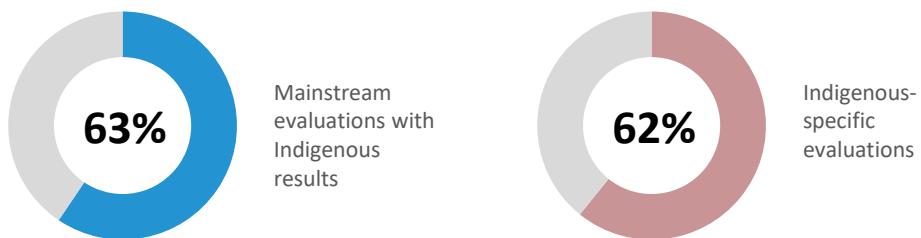
No mainstream evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people mentioned any participation of Aboriginal and Torres Strait Islander people in evaluation analysis or writing of the report (although one did have an Aboriginal organisation as part of a reference group for the evaluation), nor did they specifically provide results back to Aboriginal and Torres Strait Islander people. For Indigenous-specific evaluations just over a quarter (17) involved Aboriginal and Torres Strait Islander people in analysis or reporting and about 15 per cent (9) provided results directly back to Aboriginal and Torres Strait Islander people.

A majority of evaluations were done by external consultants

Just over 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 (figure B.20).

Information on the cost of the evaluation was only available in two of 307 evaluations provided to the Commission.

Figure B.20 Evaluations undertaken by external consultants^a



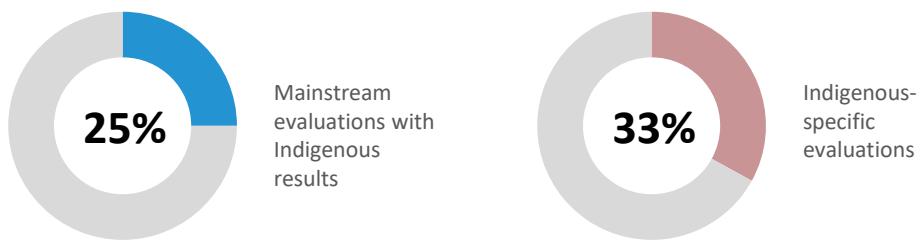
^a Proportion of evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people conducted by external consultants.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

Formal ethics assessment is uncommon

Ethics assessment was undertaken for a small proportion of evaluations — excluding evaluations where only limited information was available, one in four (14) mainstream program evaluations mentioning or providing results for Aboriginal and Torres Strait Islander people, and one in three (14) Indigenous-specific program evaluations included ethics assessment (figure B.21). Eighty per cent (22) of evaluations including ethics assessment used an established human research ethics committee. More information on agency use of ethics assessment is in figures B.13 and B.14.

Figure B.21 Evaluations that included a formal ethics assessment^a



^a Proportion of evaluations that mentioned or provided results for Aboriginal and Torres Strait Islander people that included a formal ethics assessment by a human research ethics committee. Excludes evaluations where insufficient information was available to know if ethical assessment had been undertaken.

Data source: Productivity Commission Indigenous Evaluation Strategy information request (questions 2 and 3).

There is little information on the use of evaluation results

Fewer than 10 per cent of Australian Government evaluation reports mentioning or providing results for Aboriginal and Torres Strait Islander people include information on how the Government or the commissioning agency has used the evaluation findings or recommendations. Other evaluations are probably being used but there is very little information on the usefulness of evaluation or the extent to which it assists government decision making.

B.5 Information request questionnaire



Information request to Australian Government agencies: Indigenous Evaluation Strategy

The Productivity Commission has been asked to develop a whole-of-government evaluation strategy, to be used by all Australian Government agencies, for policies and programs affecting Aboriginal and Torres Strait Islander people (including Indigenous-specific and mainstream policies and programs).

As part of this strategy, the Commission will:

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

More information about the Indigenous Evaluation Strategy is available on our website
<https://www.pc.gov.au/inquiries/current/indigenous-evaluation>.

The Commission is seeking information from Australian Government agencies on evaluation policy and activity undertaken in the last three years. Agencies' responses to this information request will inform our understanding of current evaluation arrangements and practices across the Australian Government. The information will be used to put together an aggregate picture of evaluation policy and to help us highlight examples of good practice.

We will not quote any unpublished material provided by your agency without contacting you in advance to seek your permission.

What is evaluation?

For the purposes of this survey, evaluation is defined as a systematic and objective assessment of an on-going or completed project, activity or policy. Evaluations are formal analytical endeavours involving systematic collection and analysis of qualitative and/or quantitative information.

Evaluations are undertaken to (a) improve the performance of existing interventions or policies, (b) assess their effects and impacts, and (c) inform decisions about future program design.

Source: Department of Finance, PGPA Act Glossary: <https://www.finance.gov.au/resource-management/pgpa-glossary/>.

Agency name:

Contact person name:

Position:

Telephone:

Email:

<input type="text"/>	<input type="text"/>
----------------------	----------------------

1. Does your agency have any formal policies, frameworks, strategies, plans or guides for evaluation activity conducted or commissioned by the agency? [Tick box]

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES - please provide a link to or copy of any relevant documents.

2. Over the years 2016-17 to 2019-20, has your agency planned, conducted or commissioned any evaluations of policies or programs that it has developed, administered or delivered (including both Indigenous-specific and mainstream policies and programs)?

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES – Please provide a copy of or link to any published reports of evaluations conducted or commissioned by your agency between 2016-17 and 2019-20.

Evaluations of Indigenous-specific policies or programs

Name of policy or program and link (or copy attached)
<input type="text"/>

Evaluations of mainstream policies or programs

Name of policy or program and link (or copy attached)

3. For evaluations your agency conducted or commissioned over the years 2016-17 to 2019-20, are there evaluation reports which are not currently publicly available?

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES – how many evaluations conducted or commissioned by your agency over the following years have unpublished reports?

	2016-17	2017-18	2018-19	2019-20
Evaluations conducted or commissioned but reports not yet finalised				
Evaluation reports finalised and expect them to be published in future				
Evaluation reports finalised but do not expect them to be published				
Evaluation reports finalised but unsure if they will be published or not				

If possible, please provide a copy (in confidence) of any unpublished evaluation reports

4. Does your agency have a central evaluation unit? [Tick box]

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES – what is the role of the central evaluation unit? [Tick all boxes that apply]

Conducting program or policy evaluations	<input type="checkbox"/>
Conducting cross-cutting or meta evaluations	<input type="checkbox"/>
Advising agency staff on the conduct of evaluations	<input type="checkbox"/>
Procurement of evaluation services from external providers	<input type="checkbox"/>

Providing advice to agency staff on procurement of evaluation services from external providers	<input type="checkbox"/>
Providing training to agency staff on evaluation	<input type="checkbox"/>
Setting priorities for agency-wide evaluation efforts	<input type="checkbox"/>
Collating evaluation evidence across the agency	<input type="checkbox"/>
Reviewing evaluations within the agency	<input type="checkbox"/>
Other (please specify)	

5. Does your agency have an agency-wide advisory or reference group to provide advice, oversight, governance and/or review of evaluations conducted or commissioned by the agency? [Tick box]

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES – what is the composition of the advisory or reference group? [Tick all that apply]

Senior Executives from your agency	<input type="checkbox"/>
Staff from the central evaluation unit	<input type="checkbox"/>
Staff from elsewhere in the agency	<input type="checkbox"/>
External experts	<input type="checkbox"/>
Representatives from service providers	<input type="checkbox"/>
Representatives from service user groups	<input type="checkbox"/>
Other (please specify)	

IF YES – please describe the main roles of the group.

--

IF YES – Are any members of your agency's advisory or reference group of Aboriginal and/or Torres Strait Islander origin?

Yes	<input type="radio"/>
No	<input type="radio"/>
Don't know	<input type="radio"/>

6. What criteria does your agency use to decide which policies or programs are evaluated?

7. Who within your agency decides which policies or programs are evaluated?

8. How often does your agency identify and prioritise the policies or programs that will be evaluated (e.g. annually, every two years, on an ad hoc basis, etc.)?

9. When new policies or programs are developed by your agency, an evaluation plan is: [Tick one box in each column]

Always <u>required</u>	<input type="radio"/>	Always <u>developed</u>	<input type="radio"/>
Very often <u>required</u>	<input type="radio"/>	Very often <u>developed</u>	<input type="radio"/>
Sometimes <u>required</u>	<input type="radio"/>	Sometimes <u>developed</u>	<input type="radio"/>
Rarely <u>required</u>	<input type="radio"/>	Rarely <u>developed</u>	<input type="radio"/>
Never <u>required</u>	<input type="radio"/>	Never <u>developed</u>	<input type="radio"/>
N/A	<input type="radio"/>	N/A	<input type="radio"/>

Please provide further details if desired.

10. How often is formal ethical review (e.g. approval by an ethics committee) sought for evaluation activities conducted or commissioned by your agency? [Tick box]

Always	Very often	Sometimes	Rarely	Never	N/A
<input type="radio"/>					

11. How are evaluation reports and findings shared within your agency? (For example, is there a central repository for evaluation reports, are they shared only with relevant program areas, discussed at Executive-level meetings etc.?) [free text]

12. How often are reports from evaluations conducted or commissioned by your agency made available on your agency's external website?: [Tick box]

Always	Very often	Sometimes	Rarely	Never	N/A
<input type="radio"/>					

13. Does your agency have a formal policy, strategy or guidelines for collecting and managing data, including data used in monitoring or evaluation? [Tick box]

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES - please provide a link to or copy of any relevant documents.

14. Does your agency have a formal policy, strategy or guidelines for collecting and managing information and data about Aboriginal and Torres Strait Islander people? [Tick box]

Yes	<input type="radio"/>
No	<input type="radio"/>

IF YES - please provide a link to or copy of any relevant documents.

Thank you for taking the time to complete the questionnaire

Completed questionnaires can be emailed to evaluation.survey@pc.gov.au

C A history of evaluation in the APS

This appendix provides a brief overview of the history of evaluation arrangements in the Australian Public Service (APS). It first looks at arrangements that were in place to encourage and support evaluation within public service agencies generally (section C.1). It then focuses on arrangements that were in place over time for evaluating policies and programs affecting Aboriginal and Torres Strait Islander people (section C.2).

C.1 A brief history of evaluation arrangements

Before 1988: growing recognition of the value of evaluation

Policy and program evaluation has a long history in Australia. Larger-scale efforts to promote formal evaluations for some policies and programs across the Australian Government can be traced back to at least the 1950s and 1960s (Sharp 2003, p. 6). However, the importance of evaluation in Australia began to be recognised following several significant reports in the 1970s that prosecuted the need for improved evaluation in the public service. These reports included the Coombs report of 1976 (box C.1), and the Baume report of 1979 (box C.2) (Gray and Bray 2019, p. 11; Sharp 2003, pp. 6–7).

While these reports highlighted the importance of encouraging and systematising evaluation, and the patchy state of then practice, progress establishing appropriate arrangements to improve evaluation was slow (Auditor-General 1991b, p. 12). Sharp noted that during the period of the Baume report ‘... there seemed to be an increased interest in the establishment of evaluation within government organisations’ and cited examples where evaluation did become more embedded in government decision making (2003, p. 8). But Sharp also stated that arrangements remained ‘piecemeal’, and that a consistent framework and culture for evaluation remained elusive.

Box C.1 The Coombs Report

The 1976 Royal Commission on Australian Government Administration (commonly referred to as the Coombs report after the Royal Commission's Chair) was given a broad terms of reference to 'inquire into and report upon the administrative organization and services of the Australian Government'. This included considering the:

... adequacy of the machinery available to assess the relevance and economy of existing programs in meeting government objectives ... (Royal Commission Letters Patent, point (2e))

The Coombs report made wide-ranging recommendations aimed at improving accountability and managerial practice in the public sector. Several of these recommendations centred on improving evaluation in the Australian Public Service, including that:

- there be a senior Minister for Social Welfare whose functions would include evaluating existing welfare programs (p. 429)
- a small policy unit be developed (housed within the Department of the Prime Minister and Cabinet) with part of its role to 'assist in the formulation and evaluation of programs against the government's objectives' (p. 433)
- improvements be made for collecting information for evaluating outcomes for services delivered by the states but funded by the Australian Government (p. 158, p. 421).

Source: RCAGA (1976).

One significant change from the Coombs report was the expansion of the mandate of the Auditor-General in 1979 to include the power to conduct efficiency audits. This enabled the Auditor-General to assess the economy and efficiency of departmental resource use (in addition to its existing roles that largely centred on financial auditing) (Di Francesco 1996, p. 84; McPhee 2011, p. 6). This function — in effect — continues today through the Australian National Audit Office's (ANAO) performance auditing work (ANAO 2017a).

In 1983, the newly elected Hawke Government implemented significant reforms to the administrative and budgetary processes of the APS. These reforms — collectively known as the Financial Management Improvement Program (FMIP) — aimed to enhance processes around identifying and setting budget priorities, put more focus on the goals and objectives of particular programs, develop program management techniques, and establish machinery for regular reviews of program efficiency and effectiveness (OECD 1987, p. 46).

Box C.2 The Baume Report

In 1976, the Senate Standing Committee on Social Welfare was tasked with evaluating the adequacy of Australian health and welfare services, including with reference to mechanisms to evaluate the effectiveness and efficiency of services. The Committee's subsequent 1979 report — *Through a Glass, Darkly* (commonly referred to as the Baume report after the Committee's Chair) — was largely condemning of the existing state of evaluation.

The essential prerequisites for evaluation ... are at present given low priority. The current level of activity relating to them is totally inadequate for ongoing evaluation. The evidence shows that too little ongoing evaluation is being conducted. Most of the evaluation that is undertaken consists of one-off, ad hoc studies which are inadequate for effective decision making. (p. 15)

The report espoused the need for — and the benefits of — evaluation in social and health policy and outlined the risks of not addressing the deficiencies in how evaluation was undertaken. It also explored why evaluation had been unsatisfactory, with significant findings including that evaluation was often viewed as threatening, that funding models did not encourage evaluation to be undertaken and that data necessary for evaluation did not exist or was not accessible.

The Baume report made 35 recommendations, many of which sought to improve evaluation in the Australian Public Service. The recommendations included:

- ... [that] the Department of the Prime Minister and Cabinet be charged with the responsibility for ensuring that adequate program evaluation is carried out by all Commonwealth authorities, and that it certify annually that the results of such evaluations have been properly reported to the Parliament (p. xi)
- ... [that] the Social Welfare Policy Secretariat ... prepare a document, or number of documents, outlining the methods available to organisations for the evaluation of their activities (p. xiii)
- ... [that] the Departments of Health and Social Security provide a consultancy service, free of charge, to enable organisations receiving health and welfare grants from the Federal Government to evaluate their own activities (p. xiv)
- ... [that], in future, Commonwealth funding for any health or welfare organisation be contingent on a written agreement by the organisation that it will conduct ongoing evaluation of a quality that is approved by both the organisation and the Government; and State Governments be encouraged to follow a similar practice. (p. xiv)

However, the report did stop short of recommending a centralised, systematic approach to undertaking evaluation:

It is clear to this Committee that the evaluation process ought to be the responsibility of the department providing a program or service, perhaps with access to specialised units in the Commonwealth Public Service to which it can turn for help with specific evaluation exercises. Centralised authorities should have the task of examining such evaluations with a view to ensuring that they are adequately carried out and that there is an analysis of the overall policy ramifications ... Centralised, systematic evaluation ... is neither practical nor possible. It is not practical because it does not sufficiently involve the operative staff, and thus is ineffective and more intrusive and threatening than necessary. It is not possible, because the resources that would be required would be vast and remote from the sources of data. (p. 46)

Source: Senate Standing Committee on Social Welfare (1979).

In 1987, the Australian Government issued a requirement that all New Policy Proposals (NPPs) had to include proposed arrangements for their future evaluation. Departments were also required to prepare plans for monitoring and evaluating their programs (Mackay 2011a, p. 2). A year later, the Government undertook a review of evaluation, which found significant issues with the way evaluation was being undertaken, including that:

- evaluations were not integrated into the corporate and financial decision making of agencies
- evaluations were often focused on efficiency and process, rather than the more fundamental question of overall program effectiveness
- there was a poor level of evaluation skills and analytical capacity in agencies
- the role of central departments — especially the Department of Finance (DoF) — in evaluation was unclear (Mackay 2011a, p. 3).

The Government concluded that a more centralised and prescriptive approach to evaluation was required.

1988–1996: a more formalised approach was adopted

In 1988, a Commonwealth evaluation strategy — the first whole-of-government evaluation framework at the Commonwealth level — was endorsed by the Australian Government. The strategy had three main objectives.

1. To provide information about program performance and aid Cabinet decision making and prioritisation, particularly with respect to the budget process.
2. To encourage program managers within departments to use evaluation to improve their programs.
3. To strengthen accountability by providing evidence of the performance of program managers and the management of program resources (Mackay 2011a, p. 3).

In pursuit of these objectives, the evaluation strategy placed four formal requirements on departments. These were that:

1. every program had to be evaluated every three to five years
2. each portfolio had to prepare an annual portfolio evaluation plan (outlining evaluations to occur over a three-year period), and submit these plans to the DoF
3. NPPs had to include a statement outlining arrangements for how the policy would be evaluated in the future
4. completed evaluations should be published, unless there were important policy sensitivity, national security or commercial-in-confidence considerations and that the budget documentation of departments should report major evaluation findings (Mackay 1998, p. 4).

Under the strategy, line departments were to conduct evaluations. But the DoF was given elements of an oversight role. It was empowered to provide input into agency evaluation plans and into the terms of reference of individual evaluations to ensure they aligned with government policies and priorities. The DoF was also available to participate in evaluations, subject to negotiation with the relevant line department (Mackay 2011a, pp. 3–4).

To promote compliance with the strategy, the DoF used a range of approaches. It provided departments with advice and support on evaluation, and possible access to resource agreements for line departments to aid in the management of programs found to be underperforming. The DoF's ability to influence the budget allocations of line departments was another incentive to comply with the strategy. It also produced comparative rankings of departments' approaches to planning and conducting evaluations. These rankings were never publicly released, but were shared with department secretaries. The secretary and senior executives of the DoF also advocated for evaluation, and there was explicit support for evaluation among some ministers (Mackay 2011a, p. 30).

Implementing the strategy was not straightforward. Line departments viewed the DoF's oversight as intrusive, and expressed serious concerns about its role. The strategy also necessitated a cultural change within the DoF, away from a more traditional accounting focus to one more concerned with the performance of government programs, and with greater emphasis on research and analysis (Mackay 2011a, pp. 4–5).

In time — and after recommendations from the Auditor-General (1991b, p. 47) and the House of Representatives Standing Committee on Finance and Public Administration (1990, p. 86) that the DoF be more active in monitoring and supporting departments in implementing the evaluation strategy — the DoF created a branch within its structure ‘responsible for the provision of evaluation advice, support, training and encouragement to other departments and also within DoF itself’ (Mackay 2011a, p. 5). The branch developed advice and handbooks on evaluation methodology, provided introductory training, shared evaluation best practice and promoted a community of evaluators within the public service (Mackay 2011a, p. 5).

There were mixed views about the performance of the strategy.

In 1991, about three years after the strategy was first adopted, the Auditor-General undertook a detailed examination of its performance. It found that, while a considerable number of evaluations were being undertaken under the strategy, and there was increasing attention to systematic evaluation planning, implementation of the strategy was generally poor:

... there has been a significant body of recent evidence indicating that there is still a considerable way to go before progress could be considered in anyway satisfactory ... there has been evidence of deep-seated lack of understanding and acceptance of intentions underlying the Government's evaluation strategy. Whilst evaluation is making a significant contribution to the management of change in organisations and to the sharpening of objectives, there was some broadly-based and authoritative evidence of a reluctance at many levels to address fundamental questions of effectiveness. (Auditor-General 1991b, p. 31)

Some of the recommendations stemming from the Auditor-General's assessment were directed at departments, including that central management was actively involved in implementing the strategy and reviewing evaluations to ensure that they were high quality. But several recommendations centred on the DoF playing a more active role, including by: more tightly monitoring and reporting on the performance of evaluation across the APS; being more active in providing support to agencies to develop their evaluation capacity; assuming responsibility for coordinating the preparation and dissemination of guidance material on evaluation; and coordinating the development of evaluation standards (Auditor-General 1991b).

A second Auditor-General report, focusing on the use of evaluation in preparing the budget, found that while there were notable instances where evaluation informed budget deliberations:

The level of achievement in strengthening the use of evaluation in the Budget has so far been modest. (Auditor-General 1991a, p. vii)

It found that certain aspects of the Government's framework were not being followed — for example, only one-quarter of submissions to the 1990-91 budget met the requirement to outline evaluation strategies for new program proposals — and stated that it was vital that the DoF accept an 'oversighting responsibility for monitoring, reporting, and insisting on adherence to the framework by line agencies' (p. vii).

In 1993, the Task Force on Management Improvement further examined evaluation in the public service. It found that evaluation was being incorporated into the management of the APS and that 'many agencies [were] able to demonstrate that evaluation has influenced strategic decision-making' (TFMI 1993, p. 379). However, the taskforce also found that most members of the senior executive service only 'sometimes or infrequently' used evaluation findings in their work and that, while the evaluation strategy intended to encourage a focus on results:

... ironically, there may be a tendency still to focus on satisfying the requirement for evaluation (that is, observing the letter of the strategy rather than learning to use evaluations to improve program outcomes). (TFMI 1993, p. 379)

Over the life of the strategy, evaluation increased. By the late 1980s, all Government departments had prepared annual portfolio evaluation plans and by the mid-1990s, about 160 evaluations were underway at any given time (Mackay 2011a, p. 6). Between 1993 and 1997, over 530 evaluations were published (Mackay 2011a, p. 8). There was also some evidence that there was more use of evaluation in the budget process. In the 1990-91 budget, about 23 per cent (in dollar terms) of NPPs were estimated to have been influenced by the findings of an evaluation, but by the 1994-95 budget, this was about 77 per cent (although data for the intervening years still suggested that fewer than half of NPPs were influenced by evaluation). And in the 1994-95 budget, 65 per cent of savings options were estimated to be influenced by evaluation (Department of Finance 1994b, pp. 3-4).

However, challenges remained. For example, there remained a lack of evaluation skills in many program areas in the public service, and a lack of experience in conducting or outsourcing evaluations, which had a detrimental effect on the quality of some evaluations (Mackay 2011a, p. 8). Other problems included a possible over-reliance on consultants, challenges maintaining objectivity, and scheduling the timing of evaluations so that results could inform decision making (Amies 1995, p. 132), as well as difficulties obtaining and ring-fencing evaluation budgets, and the availability and integrity of long-run datasets (Ryan 2003, p. 7).

A 1997 Auditor-General report — completed just prior to changes being implemented to unwind significant parts of the strategy — also provides insights into the conduct of evaluation under the strategy, particularly in its later years. The report examined evaluation in four agencies in detail, reviewed the quality of a sample of completed evaluations and sought information (through a survey) on the practices of portfolio agencies across the APS (Auditor-General 1997, p. xiii).

The Auditor-General found that — under the strategy — many aspects of the management and control of evaluation were ‘operating adequately’ (p. xiii). The examined departments had developed satisfactory evaluation plans, mechanisms were in place aimed at assisting evaluation quality and that while quality varied, the ‘majority of evaluations examined did not have weaknesses such that the value of the evaluation report would be affected ...’ (p. 54). However, the Auditor-General presented some concerns about departments’ access to certain evaluation skills, and commented that:

It is possible that the mandatory requirement to evaluate all programs on a three to five-year cycle may not provide sufficient scope to allow agencies to focus on strategic priorities and issues of importance to government. (p. 20)

Even after the strategy’s abolition, there were mixed views about its effectiveness. For example, while acknowledging the strategy was ‘far from perfect’ and noting that it proved to not be sustainable, Mackay suggested the evaluation strategy did result in more informed decision making:

The Australian evaluation system lasted for a decade — from 1987 to 1997. During this time, Australia became a model of evidence-based decision-making and performance-based budgeting. By the early to mid-1990s, the evaluation findings that the system was producing were having a significant influence on the contents of departments’ policy advice and on the Cabinet’s budget decisions. Ministers frequently stated the high value that they attached to having evaluation findings available to better inform their decision making. Evaluation findings were also being used quite intensively within line departments in support of their ongoing management. (2011a, p. 27)

In contrast, Ryan, while acknowledging that the positives that emerged from the strategy should not be ignored, suggested that the strategy did try to do too much in too little time:

The 1988 Commonwealth Evaluation Strategy *did* expect too much too soon. Moreover, evaluation was regarded as a straightforward management tool that only had to be used to start producing results. The immediate effect was evaluation overload, patchy quality, unselective evaluation, and a certain waste of money. (2003, p. 8, emphasis in original)

1996–2007: a shift to more decentralisation

Just as the Hawke Government implemented a suite of reforms for the federal public service not long after entering office, so did the Howard Government when it was elected in 1996. The new Government sought to reduce bureaucratic ‘red tape’, promote a greater role for the private sector in policy design and service delivery, and make department secretaries (and the public service as a whole) more responsive to political priorities of government (Mackay 2011a, p. 11).

Line departments highlighted the burdens placed on them to plan and undertake evaluations to the new Government, and pushed for less oversight from — and reporting to — the DoF (Mackay 2011a, p. 14). In 1997, the strategy was abolished. Reasons given for the strategy’s abolition included that it was too cumbersome, that it was too resource intensive for all parties and that there remained issues with the state of program evaluation skills in the public service (Mackay 2011a, p. 14; Tune 2010).

The support that the DoF provided to line departments to undertake evaluation was discontinued. While the DoF adopted a principles-based approach to encourage departments to evaluate key policies and programs, the decision whether or not to evaluate essentially laid with individual department secretaries and the extent that agencies continued to prioritise evaluation varied (Mackay 2011a, pp. 14–15). That said, agencies were still required to review lapsing programs, although this process was described as ‘mechanical’ with little impact on decision making, and was eventually abolished (Blöndal et al. 2008, p. 19; Hawke and Wanna 2010, p. 78).

An ANAO report in 2003 found that most agencies were still undertaking evaluations, but evaluations were not frequently discussed in annual reports or used to support performance reporting (ANAO 2003, p. 13).

Further reforms came in 1999 with the introduction of the *Outcomes and Outputs Framework*. The framework was introduced to encourage agencies to ‘focus on ends and not means’ (Webb 2002, p. 12). Agencies were required to report on planned outcomes (results or consequences for the community that the Government sought to achieve) and on outputs (goods and services they produced in order to contribute to these outcomes) (Webb 2002, p. 2).

Under the framework, related programs were allowed to be aggregated ‘with the intention of giving more emphasis to the overall impact — and effectiveness — of government activity in that field’ (Podger 2018, p. 109). Bray and Gray stated this was the subject of significant debate:

The aggregation of programs into ‘outcomes’ was the subject of considerable debate over subsequent years including whether the outcomes were overly broad and not conducive to practical accountability oversight by the Parliament, and whether the connections between inputs and outcomes were too uncertain. (2019, pp. 11–12)

Podger said that there was a less systematic approach to evaluation under the framework, and it seemed that it ‘was not given the priority it had attracted previously’ (2018, p. 110).

In 2007, the ANAO conducted an audit of the application of the Outcomes and Outputs Framework. As part of this audit, it surveyed 44 agencies and undertook an in-depth audit of three agencies. It found that while ‘the majority of agencies used outcomes and outputs information in [their] decision-making’ (p. 27), the audited agencies:

... did not generally report on the effectiveness of administered items in achieving outcomes. Agencies often provided details of activities undertaken or funding provided without demonstrating how these activities achieved programme objectives and, ultimately, outcomes. (ANAO 2007, p. 80)

2007 onwards: more changes

The election of the Rudd Government in 2007 brought several significant reviews across a range of policy areas, but there was no systematic change in the use of evaluation in the budget process, nor was there a consistent approach to publishing evaluations (Gray and Bray 2019, p. 12).

In 2009, the Outcomes and Outputs Framework was replaced with the *Outcomes and Programs Framework*. A central requirement of the new framework was that entities were to identify and report against programs that contributed to government outcomes, rather than reporting on output groups, meaning that programs became the building blocks of budgeting and reporting. Annual reporting on the delivery of programs and achievement against a set of Key Performance Indicators (KPIs) was a key element of the framework (ANAO 2011b, pp. 14–15, 37).

The application of the Outcomes and Programs Framework was examined by the ANAO in both 2011 — where it found that most of the agencies examined had scope to improve their development of KPIs and the reporting against them (and in some cases, significantly so) (ANAO 2011b, p. 17) — and in 2013 — where it found that agencies continued to experience challenges in developing and implementing meaningful indicators (ANAO 2013b, p. 19). The 2013 audit also found:

The broad range of roles and activities undertaken by Australian Government entities highlights that while some entities’ programs suit the development of relatively straightforward performance information, the less tangible nature of the objectives of some other programs is more of a challenge. A homogenous framework for application by all Australian Government entities, without recognition of the variety of entity activity/ies, has compounded challenges that entities have in implementing the Australian Government performance measurement and reporting framework. (p. 59)

In 2013, further public sector governance reforms led to the passing of the *Public Governance, Performance and Accountability Act 2013* (Cwlth) (PGPA Act). The PGPA Act updated performance reporting for Commonwealth entities by combining the

requirements set out in two existing public sector governance Acts⁷¹ and placing them 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. The PGPA Act remains in operation today and the extent to which it promotes evaluation is discussed in chapter 2.

C.2 Evaluation of policies and programs affecting Aboriginal and Torres Strait Islander people

As well as recommending changes to the APS to improve evaluation more generally, the 1976 Coombs Report identified a particular 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 report recommended that the Department of Aboriginal Affairs and the Department of the Prime Minister and Cabinet periodically assess programs in the Aboriginal affairs policy area. It also recommended that the Department of Aboriginal Affairs and the Institute of Aboriginal Studies collaborate in research into possible indicators of social welfare for Aboriginal people (RCAGA 1976, pp. 430–431).

In 1979, the House of Representatives Standing Committee on Aboriginal Affairs also called for more evaluation of programs targeted at Aboriginal and Torres Strait Islander people, this time in the context of health policy:

The Committee recommends that an independent evaluation team responsible to the Minister for Aboriginal Affairs be established to evaluate the effectiveness of all Aboriginal health care services and programs in accordance with the World Health Organisation’s definition of health and the principles of self-determination ... (House of Representatives Standing Committee on Aboriginal Affairs 1979, p. 109)

At about the same time, a dedicated evaluation branch was established within the Commonwealth Department of Aboriginal Affairs, whose remit included to ‘review and evaluate the effectiveness of Government programs in Aboriginal communities’ (Department of Aboriginal Affairs 1981, p. 49). In 1982–83, this function was merged into the Department’s budgeting branch (Department of Aboriginal Affairs 1983, p. 55).

In 1990, the Hawke Government established the Aboriginal and Torres Strait Islander Commission (ATSIC), and this brought major changes to how policies and programs affecting Aboriginal and Torres Strait Islander people were evaluated.

⁷¹ These Acts were the *Commonwealth Authorities and Companies Act 1997* (Cwlth) and the *Financial Management and Accountability Act 1997* (Cwlth) (AGS 2013).

Under the *Aboriginal and Torres Strait Islander Commission Act 1989* (Cwlth), the Government also established the Office of Evaluation and Audit (OEA). The OEA was independent from the ATSIC Board of Commissioners, with a legislated remit that included evaluating and auditing the operations of ATSIC regularly (ss. 75-76). The Act also requested that the director of the OEA — in overseeing the operations of the Office — have regard to the desirability of evaluating and auditing programs at least once every three years (s. 78). Between 1990 and 2005, the Office evaluated a range of programs, including programs related to Native Title, housing, health, community infrastructure and legal services.

In 2005, ATSIC was abolished, but the OEA was retained under revised legislation — the *Aboriginal and Torres Strait Islander Act 2005* (Cwlth) — and moved to the then Department of Finance and Deregulation (Plowman 2008, p. 2). Its name was changed to the Office of Evaluation and Audit (Indigenous Programs) and its role redefined to include evaluating and auditing relevant programs administered by Australian Government bodies (s. 193W). In 2009, the Office was incorporated into the ANAO (in effect, ceasing as a standalone office) (ANAO 2011a, p. 1).

A coordinating office — the Office of Indigenous Policy Coordination (OIPC) — was also established within the then Department of Immigration and Multicultural and Indigenous Affairs in 2004. Its functions included ‘being the primary source of advice on Indigenous issues to the Government, coordinating and driving whole-of-government innovative policy development and service delivery ... and overseeing relations with state and territory governments on Indigenous issues’ (ANAO 2012, pp. 14–15). The Office’s role also included evaluating and reporting on the performance of government programs and services for Aboriginal and Torres Strait Islander people (ANAO 2012, p. 35). In 2006, the OIPC’s functions were absorbed into the Department of Families, Community Services and Indigenous Affairs (ANAO 2012, p. 15).

In 2007, the Council of Australian Governments (COAG) agreed to a partnership ‘to achieve the target of Closing the Gap in Indigenous disadvantage’ (COAG 2009, p. 3) and in 2009, the Closing the Gap Clearinghouse was established. Its key objective was to:

... improve access to evidence by analysing research and evaluation findings to identify what works to address Indigenous disadvantage. (ANAO 2019a, p. 57)

The Clearinghouse was jointly funded by the Australian, state and territory governments and jointly delivered by the Australian Institute of Health and Welfare and the Australian Institute of Family Studies. It maintained a register of government research and evaluations (which, by March 2014, contained over 1000 items) and produced 50 evidence synthesis reports (ANAO 2019a, p. 57). It also produced three ‘what works’ reports. A lack of robust evaluation of programs was a consistent theme in these reports:

A common issue identified in the ‘what works’ reports was a lack of robust evaluations to inform program development. In particular, the reports noted there was a lack of high-quality quantitative social policy research in the Australian and Aboriginal and Torres Strait Islander contexts. (ANAO 2019a, p. 57)

A 2010 strategic review of Indigenous expenditure also commented on the state of evaluation of policies and programs targeted at Aboriginal and Torres Strait Islander people. It described evaluation activity as patchy and noted:

... the poor quality and lack of rigour of many of the evaluations that have been undertaken. Serious data limitations aside, many evaluations appear to have been conducted either perfunctorily — simply because they are required to be done — or for self-serving reasons, with a view to perpetuating the program. (DoFD 2010, p. 366)

In 2014, funding arrangements for the Clearinghouse were set to expire. The Department of the Prime Minister and Cabinet advised that there was insufficient new research and evaluation being produced to sustain the existing rate of publication of the Clearinghouse. The Minister for Indigenous Affairs gave in principle agreement to fund the Clearinghouse at a reduced scale for an additional three years, pending the state and territory governments matching the Commonwealth's contribution. However, the Australian Government could not secure an agreement from all jurisdictions and funding for the Clearinghouse ceased (ANAO 2019a, pp. 57–58).

With the closure of the Clearinghouse, there was no whole-of-government approach to publishing and analysing evaluation reports for policies and programs affecting Aboriginal and Torres Strait Islander people. That said, evaluation continued to occur at an agency/departmental level and the ANAO continued to conduct performance audits of policies and programs affecting Aboriginal and Torres Strait Islander people (including in the areas of Indigenous childhood development (2013a), service delivery (2014a), legal assistance (2015) and aged care (2017c)).

In 2013, the Indigenous Advancement Strategy was established, which saw responsibility for most programs for Aboriginal and Torres Strait Islander people (as well as some mainstream programs that primarily serviced Aboriginal and Torres Strait Islander people) transferred to the Department of the Prime Minister and Cabinet (ANAO 2017b, p. 15). In February 2018, the Department released an evaluation framework document (ANAO 2019b, p. 7). The implementation of — and the evaluation of programs under — the Indigenous Advancement Strategy is undertaken by the National Indigenous Australians Agency today.

D International approaches to evaluation

This appendix provides a high-level overview of how governments of other countries approach evaluation — both in a broad sense, and specifically for policies and programs affecting indigenous peoples. The appendix highlights some key aspects of other countries' approaches to evaluation.

The appendix first takes a cross-country approach to examine:

- the use of central evaluation frameworks in other countries (section D.1)
- how other countries seek to ensure quality in their evaluations (section D.2)
- how other countries promote evaluation use in their decision making (section D.3).

It then looks at the way Canada, the United States and New Zealand approach evaluating policies and programs that affect indigenous peoples (section D.4).

D.1 Central evaluation frameworks to govern evaluation

Central frameworks seek to institutionalise evaluation by formally establishing arrangements about how evaluation will be organised and governed on a whole-of-government level. In some cases, these frameworks are established through law — for example, in **France**, a legal framework for evaluation is embedded in primary and secondary legislation, as well as in the French Constitution (OECD 2020b, p. 40).

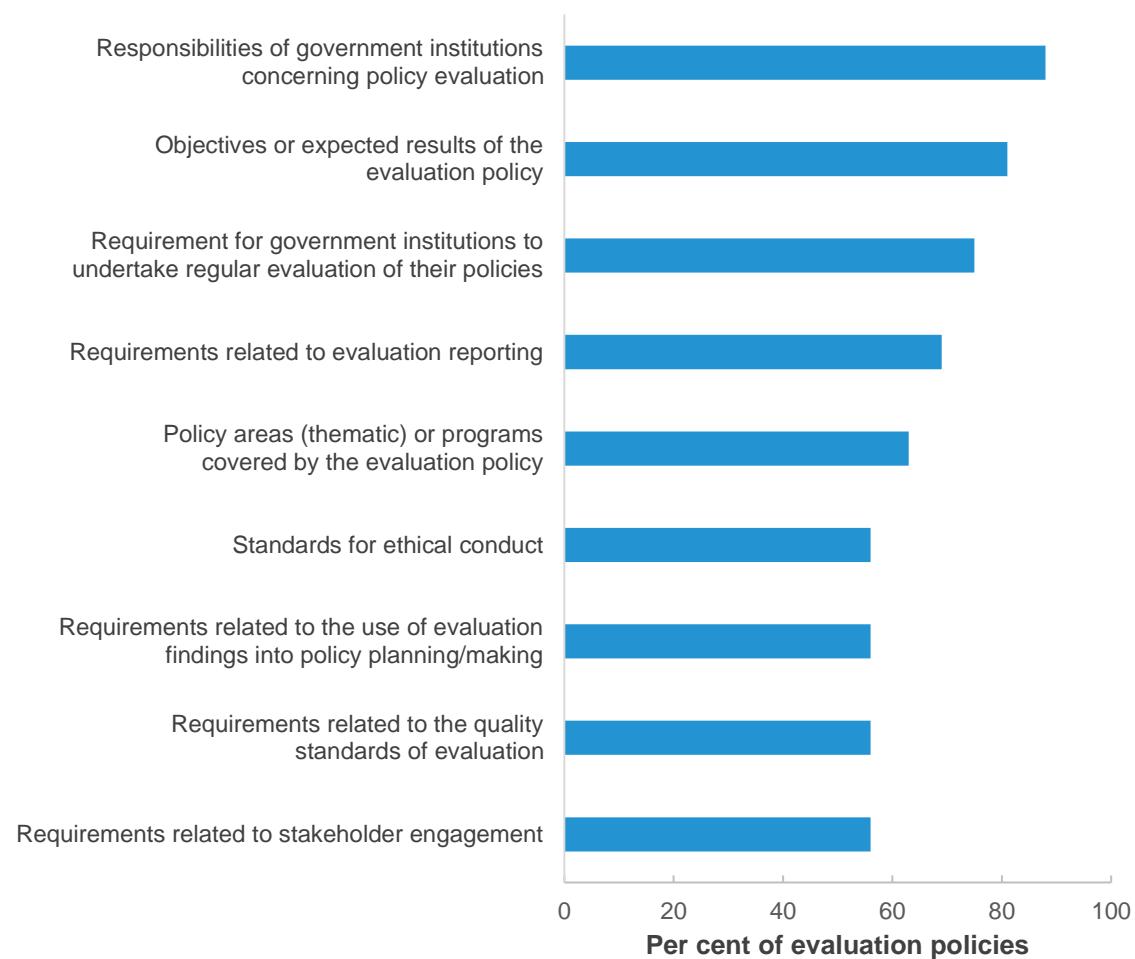
Central evaluation frameworks can also be established through policy (with or without legislative underpinnings). Evaluation policies are reasonably common. A recent survey of 42 countries — including all Organisation of Economic Cooperation and Development (OECD) member countries with the exception of Luxembourg — found that half have developed a policy framework to organise evaluation activities across government (OECD 2020b, p. 44).

There are a number of common elements in the evaluation frameworks of different countries. For example, of the frameworks developed by OECD countries:

- 88 per cent outlined the responsibilities of government institutions concerning evaluation
- 81 per cent outlined the objectives or expected results of the evaluation policy
- 75 per cent contained requirements for government institutions to undertake regular evaluations of their policies (figure D.1).

Figure D.1 There are common features across the evaluation policies of different countries

Per cent of evaluation policies that contain the identified element (OECD member countries only)



Data source: OECD (2020b, p. 45).

Canada's Policy on Results provides a useful example of an evaluation policy that has several of these elements. The Policy on Results was implemented in 2016, and represents the latest iteration of a long line of central evaluation strategies put in place by the Canadian Federal Government from the late 1970s (TBS 2010).

The Policy on Results has two overall objectives. First, it seeks to improve the achievement of results across government, and second, it seeks to enhance the understanding of government goals, achievements and resourcing inputs (Government of Canada 2016b). These objectives are pursued, in part, by placing a number of obligations on federal departments for performance measurement and evaluation. These obligations include:

- establishing a *Departmental Results Framework* that sets out:
 - the core responsibilities of the department
 - departmental results (defined as the changes departments seek to influence)
 - departmental result indicators (factors or variables that provide a valid and reliable means to measure or describe progress on a departmental result)
- implementing and maintaining a *Program Inventory* that identifies all of the department's programs and how resources are organised to contribute to the department's responsibilities and results
- developing *Performance Information Profiles* that identify performance information for each program
- developing, implementing and releasing a *rolling five-year evaluation plan*. This plan must include evaluations — at least once every five years — of all ongoing programs of grants and contributions with a five-year average (actual) expenditure of CAD \$5 million or more per year⁷² (Government of Canada 2016a, 2016b).

The Policy on Results also requires each department to establish several roles to oversee evaluation. This includes establishing a *Performance Measurement and Evaluation Committee* — chaired by deputy department heads — to oversee departmental performance measurement and evaluation, and a *Head of Evaluation* who is responsible for leading evaluation within the department. The department must also designate a *Program Official* for each of its programs who is responsible for — among other functions — ensuring that valid, reliable and useful performance information is collected for that program (figure D.2) (Government of Canada 2016b, 2016a).

Implementation of — and compliance with — the Policy on Results is overseen by the Canadian Treasury Board Secretariat.⁷³ Its 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 deemed necessary

⁷² s. 42.1 of the *Financial Administration Act 1985* (Canada) also sets out a requirement that ongoing programs be reviewed.

⁷³ The Treasury Board Secretariat is a central government agency (TBS 2018). It provides advice to the Treasury Board (a committee of Cabinet Ministers) on 'how the government spends money on programs and services, how it regulates and how it is managed' (Government of Canada 2020e).

- initiating centrally-led evaluations if deemed necessary
- raising any issues with compliance with the Policy with deputy heads or the President of the Treasury Board (a Minister)
- establishing competencies for the heads of performance measurement and heads of evaluation and amending these competencies as appropriate (Government of Canada 2016b).

Figure D.2 Roles required under the Policy on Results

	Department Deputy Head Ensures adherence to the Policy on Results
	Performance Measurement and Evaluation Committee Oversees departmental performance measurement and evaluation
	Head of Performance Measurement Leads the departmental performance measurement function
	Head of Evaluation Leads the departmental evaluation function
	Program Officials Maintain performance information for their program

Source: Adapted from Pagan (2016).

When undertaking evaluations, the Directive accompanying the Policy on Results also mandates that Canadian departments follow a broad set of requirements called the *Standard on Evaluation* (Government of Canada 2016a). This includes that evaluations are planned with consideration of the ‘primary evaluation issues’ of relevance, effectiveness and efficiency where relevant to the goals of the evaluation (box D.1).

The Standard on Evaluation also stipulates some requirements for evaluation reports, including that they:

- be written and presented clearly and concisely
- include the information required to understand and reasonably sustain findings and conclusions, and present a logical flow of findings, conclusions and recommendations
- provide readers with an appropriate context for the evaluation and the policy, program, priority, unit or theme being evaluated
- identify the limitations of the evaluation in a way that informs readers about the reliability of findings and conclusions

-
- include an accurate assessment of the contribution of the program to its related government priorities and/or departmental results and priorities
 - include clear, actionable recommendations that aim to address the key issues or concerns identified (Government of Canada 2016a).

Box D.1 **Relevance, effectiveness and efficiency as defined by the Canadian Standard on Evaluation**

The Standard on Evaluation defines **relevance** as the extent to which a policy or program addresses and is responsive to a demonstrable need. Relevance may also consider if a program or policy is a government priority or federal responsibility.

It defines **effectiveness** as the impacts of a policy or program, or the extent to which it is achieving its expected outcomes.

It defines **efficiency** as the extent to which resources are used such that a greater level of output/outcome is produced with the same level of input, or a lower level of input is used to produce the same output/outcome.

Source: Government of Canada (2016a).

The *Foundations for Evidence-Based Policymaking Act 2018* (US) of the **United States** is another example of a whole-of-government framework for evaluation. The Act seeks to advance data and evidence building functions in the US Federal Government (Executive Office of the President (US) 2020b s. 290.3). Similar to the Canadian Policy on Results, the Act places several formal obligations on US federal agencies to plan and undertake evaluation. One of these obligations requires agencies to develop an *evidence building plan*. An evidence building plan ‘is a systematic plan for identifying and addressing policy questions relevant to the programs, policies and regulations of the agency’ that ‘identifies, prioritizes, and establishes strategies to develop evidence to answer important short- and long-term strategic questions ... and operational questions ... ’ (Executive Office of the President (US) 2020b s. 290.7).

Under the Act, agencies are required to list (among other things) in their plans:

- policy-relevant questions for which the agency intends to develop evidence to support policymaking
- data that the agency intends to collect, use or acquire to facilitate the use of evidence in policymaking
- methods and analytical approaches that may be used to develop evidence to support policymaking
- any challenges to developing evidence to support policymaking, including any statutory or other restrictions to accessing relevant data.⁷⁴

⁷⁴ s. 312 of the *Foundations for Evidence-Based Policymaking Act 2018* (US) outlines these requirements.

The Act also requires agencies to *issue an evaluation plan* for each fiscal year that describes the evaluation activities it plans to conduct.⁷⁵ These evaluation plans are required to describe:

- the key questions that are to be addressed by each significant evaluation. This may include discussion of a program’s purpose, goals and objectives and how program activities are linked to their intended effects
- the information needed for evaluations, including whether new information will be collected or whether existing information will be acquired
- 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) 2019, pp. 34–35).

The Act also obliges agency heads to designate a senior employee as the agency’s *evaluation officer*.⁷⁶ Evaluation officers have substantial and wide-ranging responsibilities — from the broad (such as being a champion for evaluation in the agency), to the more specific (such as overseeing the establishment and implementation of an agency evaluation policy) (box D.2).

The Act requires the Office of Management and Budget to issue guidance for program evaluation for agencies that is consistent with widely accepted standards for evaluation and to identify best practices.⁷⁷ This guidance was released in March 2020 (Executive Office of the President (US) 2020a, p. 2).

D.2 Approaches to promoting evaluation quality

Governments (in countries both with and without central evaluation frameworks) have adopted a range of approaches to encourage agencies to produce high-quality evaluations. These approaches include (but are not limited to):

- developing evaluation guidelines
- developing and promoting evaluator competencies
- using review mechanisms to promote evaluation quality (OECD 2020b, p. 89).

⁷⁵ s. 312 (b).

⁷⁶ s. 313.

⁷⁷ s. 315 (e)(1).

Box D.2 **Responsibilities of evaluation officers in the United States**

The expected responsibilities of evaluation officers are spelled out in a circular sent from the Executive Office of the President regarding the preparation, submission and execution of the budget in July 2020. These responsibilities include (but are not limited to):

- serving as an agency champion for — and educator of — agency staff and leaders about evaluation, including what evaluation is, its value, how to discern high-quality evaluation from other types of analysis, and the importance of evaluation as a strategic investment
- serving as a senior advisor to agency leaders on issues of evaluation policy and practice
- serving as a senior agency contact on evaluation for agency-wide and cross-cutting evaluation efforts.

Evaluation officers also oversee or conduct:

- assessments of the coverage, quality, methods, effectiveness, objectivity, scientific integrity and balance of the portfolio of evaluations, policy research and ongoing evaluation activities of their agency
- improvements of agency capacity to support the development and use of evaluation
- the establishment and implementation of an agency evaluation policy that affirms the agency's commitment to conducting rigorous, relevant evaluations and to using evidence from evaluations to inform policy and practice
- the required coordination, development and implementation of plans required under the Act, including annual evaluation plans and evidence-building plans
- the development of new or improved processes to integrate evaluation findings into agency decision making and other functions
- management of agency-wide evaluation policies that uphold and adhere to identified standards
- the use and dissemination of evaluation results throughout the agency and to the public, as appropriate.

Source: Executive Office of the President (2020b s. 290.5).

Developing evaluation guidelines

Evaluation guidelines are common — according to a recent survey, about three-quarters of OECD member countries have evaluation guidelines in place (OECD 2020b, p. 90). The guidelines contain a range of features designed to promote high-quality evaluations (figure D.3). Some of these features relate to technical aspects of evaluation (such as guidance on quality standards or evaluation design) while others centre on promoting good evaluation governance arrangements (such as guidance on how to promote evaluation independence).

Figure D.3 Evaluation guidelines provide guidance on a wide range of matters

Per cent of evaluation guidelines that contain guidance on the identified element (OECD member countries only)



Data source: OECD (2020b, p. 92).

The Magenta Book, published by the **United Kingdom (UK)** Treasury Department (HM Treasury), for example, provides guidance on evaluation for UK government departments (HM Treasury 2020c). The Magenta Book is one of a series of guides produced by HM Treasury to better incorporate evidence into policymaking — other resources include the Green Book, which assists agencies to appraise policy options (with a focus on cost-benefit analysis), and the Aqua Book, which provides guidance on producing (and commissioning) high-quality analysis to inform policymaking (HM Treasury 2015, 2018).

The Magenta Book provides guidance on: evaluation scoping; methods; data; managing evaluations; the use and dissemination of findings; and capabilities that evaluation managers should have. It also outlines some principles that can be used to guide decision making to maintain high-quality evaluation (box D.3).

Box D.3 Evaluation principles set out in the Magenta Book

The Magenta Book outlines four principles to guide decision making with respect to evaluation. These principles are:

- **Useful** — an evaluation is of high quality when it is designed to meet the needs of stakeholders and ‘produces useful, usable outputs at the right point in time’.
- **Credible** — credibility and usefulness go hand in hand. Objectivity is often important to producing a credible evaluation and transparency is also crucial.
- **Robust** — evaluations should be well-designed, use an appropriate evaluation approach and methods, and be executed well. When measuring impact, the evaluation approach should undertake comparisons, either across time, between groups, or between alternative theories. Using peer review and independent steering is identified as a way to quality assure an evaluation’s design and execution.
- **Proportionate** — not all interventions require the same amount of scrutiny or have the same needs for learning, and at times, a light-touch monitoring and evaluation approach may be all that is needed. The Magenta Book outlines criteria for ‘priority’ interventions that require ‘substantial’ evaluations. The criteria include interventions that: are high profile policies; have high levels of uncertainty or risk; are high cost; and have high learning potential.

Source: HM Treasury (2020c, p. 16).

There are several supplementary papers to the Magenta Book. These papers provide additional guidance to policymakers on doing evaluation well in specific contexts. This includes examining quality in the context of qualitative evaluations, examining quality in impact evaluation (with a focus on empirical impact evaluations), handling complexity in evaluations, and undertaking realist evaluations (Campbell and Harper 2012; HM Treasury 2012, 2020b, 2020a).

Establishing and promoting evaluator competencies

Many OECD countries have developed mechanisms to support evaluators to develop their competencies (OECD 2020b, p. 108). This commonly takes the form of training, which is provided in many different ways. For example:

- in **Canada**, the Treasury Board Secretariat, in collaboration with the Canada School of Public Service, runs an evaluation learning series, with a focus on strategic evaluation for senior evaluators, managers and directors. Recent sessions in this series have focused on evaluation planning, evaluation conduct and strategic communication of evaluation results (Government of Canada 2020b)
- in **Slovakia**, the Slovak Value for Money Institute maintains a list of recommended courses for evaluators, and analysts who join policy evaluation units are offered opportunities to undertake a range of courses to deepen their evaluation knowledge in specific policy areas. Additionally, before entering an analytical team in a Ministry, analysts working on evaluation are required to pass a test that assesses their competencies (OECD 2020b, p. 111).

Some countries also build or promote evaluation competencies into the functional or professional streams of their public service. For example, in the **United Kingdom**, the Government Social Research Profession Stream and the Policy Profession Stream outline competencies and/or skills related to evaluation that certain staff in these streams should have (Government Social Research (UK) 2019; Policy Profession (UK) 2019).⁷⁸ In addition, an evaluation capability framework (and a self-assessment tool) has been developed for Government analysts (HM Treasury nda, ndb).

Elements of a profession-based approach to evaluation in the public service are also emerging in the **United States**, with the *Foundations for Evidence-based Policymaking Act 2018* (US) requiring the Office of Personnel Management (the chief human resources agency for the US Federal Government) to:

- identify the key skills and competencies needed for program evaluation in an agency
- establish a new occupational series for program evaluation (or update and improve an existing occupational series)
- establish a new career path for program evaluation within an agency.⁷⁹

Another mechanism used in some countries to develop and maintain evaluator competencies is government supported evaluation networks (box D.4). These networks sit aside or complement evaluation networks that operate outside government, such as those facilitated through national evaluation associations (which exist in all OECD member countries) (OECD 2020b, p. 117).

In some countries, there are processes to support evaluators with additional resourcing and expertise to help them undertake evaluations. For example, in the **United States**, the Office of Evaluation Sciences (a government office with a team of interdisciplinary experts) partners with US federal agencies to help them build and use evidence, including by assisting agencies to undertake randomised evaluations (Office of Evaluation Sciences (US) nda, ndb). The Office has also developed toolkits for US agencies to aid them in meeting their obligations under the *Foundations for Evidence-based Policymaking Act 2018* (US) (Office of Evaluation Sciences (US) 2020).

Review mechanisms to promote evaluation quality

Compared with other levers to influence evaluation quality (such as those discussed above), the use of review mechanisms to ensure evaluations are of a high standard appears less common.

⁷⁸ The Government Social Research Profession and the Policy Profession are two of the 28 recognised professional streams in the UK Civil Service. Each profession is led by a Head of Profession (Government of the United Kingdom nd).

⁷⁹ s. 315.

Box D.4 Some examples of government supported evaluation networks

The Evaluation Officer Council (United States)

The Evaluation Officer Council is in the process of being established. The Council — to be made up of agency Evaluation Officers — will serve as a forum to: exchange information; consult with and advise the Office of Management and Budget (which is overseeing the implementation of the *Foundations for Evidence-based Policymaking Act 2018* (US)) on matters that affect evaluation; coordinate and collaborate on common interests; and to provide a leadership role for the wider Federal evaluation community.

The EVA-forum (Norway)

The EVA-forum is an informal network of evaluators chaired by a representative of the Norwegian Agency for Financial Management. The forum provides members with an opportunity to share their experiences on evaluation, including by organising several networking and workshop events per year, as well as an annual national evaluation conference.

Sources: Executive Office of the President (US) (2019, p. 10); OECD (2020b, pp. 113–114).

While there are examples of peer review processes being used to promote evaluation quality (including for particular evaluations in **Portugal** and **Germany**) (OECD 2020b, p. 105), requirements to formally review evaluations for quality control do not appear to be particularly widespread. (Some evaluation guidelines, such as the **United Kingdom's** Magenta Book and **Canada's** Standard on Evaluation promote peer review as a way to provide quality assurance for evaluation, but the extent that peer review is used — even when encouraged — remains unclear.)

In **Japan**, the Administrative Evaluation Bureau (which sits within the Ministry of Internal Affairs and Communications) has a form of whole-of-government review function. The Bureau reviews policy evaluations that ministries have conducted and identifies elements in these evaluations that need improvement⁸⁰ (OECD 2020a, p. 59). A broad government-wide review function had also existed in **Canada**, where the Treasury Board Secretariat published an annual report on the *Health of Evaluation Function*. These reports provided a snapshot of evaluation in departments and agencies, and included an assessment of departments' performance in undertaking quality evaluation. However, this annual reporting appears to have ceased — the most recent Health of Evaluation Function report that the Commission was able to locate is for 2013 (Centre of Evaluation Excellence 2015).

Some countries have developed self-evaluation tools for evaluators to review and assess the quality of evaluations but this practice is not particularly widespread (with responses to a recent survey suggesting only two OECD countries — **Poland** and **Spain** — have used self-evaluation checklists) (OECD 2020b, p. 108).

⁸⁰ The Bureau also carries out cross-ministry evaluations and formulates rules and guidelines for conducting policy evaluations (OECD 2020b, p. 42, 2020a, p. 59).

D.3 Approaches to promoting evaluation use

Evaluation is often undertaken with the intent to generate better evidence to inform decisions about how policies and programs should be designed, structured and funded. Similar to how governments have taken steps to promote high-quality evaluation, there are also levers available to governments to promote the use of evaluation in policymaking and decision-making processes. These levers include:

- encouraging evaluation publication and translation
- establishing evaluation and evidence champion positions or functions within and across government agencies
- requiring management responses to evaluation.

Encouraging evaluation publication and translation

Just under half of OECD member countries make evaluation findings and recommendations publicly available by default (OECD 2020b, p. 127). This is done in various ways. For example, in **Canada**, the Policy on Results stipulates that evaluation reports and summaries are released on web platforms (as prescribed by the Treasury Board Secretariat), and evaluation reports are published in a specific section of the website of the department or agency that they relate to (Government of Canada 2016b, 2019). In contrast, in **Norway**, evaluations carried out on behalf of government agencies are published in a single publicly-accessible online evaluation repository, jointly managed by the Norwegian Directorate for Financial Management and the National Library of Norway (OECD 2020b, p. 128).

Some countries have also established approaches for translating evaluation knowledge. Defined broadly, these approaches seek to enhance the usability and usefulness of evaluation findings by contextualising and/or synthesising the findings to produce clear, actionable and policy-relevant evidence.

The *What Works Network* in the **United Kingdom** provides an example of knowledge translation in practice. Established in 2013, the network currently consists of nine centres (with an additional three affiliate members and one associate member) (Government of the United Kingdom 2019). Each centre focuses on a unique policy area⁸¹ and operates in their own way, but all centres share a common remit that includes knowledge translation (box D.5).

⁸¹ These policy areas are: health and social care; educational achievement; crime reduction; early intervention; local economic growth; improved quality of life for older people; wellbeing; homelessness; children's social care; youth offending; youth employment; and higher education.

Box D.5 The remit of What Works Centres

The overall objective of the What Works Network is to:

... provide the best evidence of ‘what works’ to the people who make decisions about public services. These people include, among others, government ministers, civil servants, council leaders, service commissioners, citizens themselves and professionals from across the public sector. (p. 1)

To pursue this objective, What Works Centres are required to undertake three core functions. The first of these functions focuses on generating high quality and relevant evidence on what works — and what does not — for their specified policy area. This includes identifying research gaps and working with research partners to fill them. Another function centres on improving the use of and demand for high-quality evidence among decision makers. The final function relates to knowledge translation.

On knowledge translation, What Works Centres:

... identify decision-makers in their policy area, and commit to translating technical research into a format they can understand and use. They strive to understand their users’ needs and put them at the heart of everything they do. (p. 5)

In delivering this translation function, What Works Centres are required to:

- publicise and share evidence generated with users, providing guidance on how to interpret and use the information and adapting style where needed
- maintain independence and methodological rigour when communicating the evidence, making sure that advice to users does not reach beyond what the evidence suggests and primacy is given to findings from high-quality impact evaluations
- commit to a high level of transparency around all evidence and recommendations to decision-makers
- use plain English explanations of the limitations of any advice to ensure that users do not misuse the evidence published by the centre. (p. 5)

Source: Cabinet Office (UK) (2018).

What Works Centres are funded through a combination of government and non-government contributions, and different centres have different levels of funding (Government of the United Kingdom 2019; What Works Network 2018). Membership to the What Works Network requires a commitment to uphold a number of principles, including independence (which includes retaining editorial control over all research and products), methodological rigour, capacity building and transparency (Cabinet Office (UK) 2018, p. 3).

In their first five years (from 2013 to 2018), What Works Centres have collectively produced or commissioned nearly 50 systematic reviews of the evidence base within their policy areas, and five centres have produced evidence comparison toolkits that use common measures to compare interventions on the basis of their impact, cost and the strength of the evidence supporting them (What Works Network 2018, pp. 11, 21).

Use of evaluation and evidence champions

Several countries have established functions — of varying degrees of separation from government — to ‘champion’ or promote the effective use of evaluation (and/or evidence more broadly) within government agencies. An evaluation or evidence champion function may be

an individual or an organisation, or be spread across a network. This role often sits within a wider remit, as is the case with the **United Kingdom's** What Works Centres, which are required to ‘commit to improving the use of and demand for high-quality evidence among decision makers’ in addition to their knowledge generation and translation functions (discussed above) (Cabinet Office (UK) 2018, p. 5). As part of this commitment, centres are tasked with providing advice and encouragement for those commissioning and delivering innovative interventions to evaluate their work, and building capacity within user groups to understand, use and build evidence.

In the **United States**, one of the core functions of evaluation officers is to serve as a champion for evaluation within their agency (box D.2).

In **New Zealand**, a form of championing function appears to reside with the Prime Minister’s Chief Scientific Advisor (PMCSA). While the role of the PMCSA is to provide high-quality, independent, scientific advice to the Prime Minister and the Government, the terms of reference for the PMCSA note that the Advisor may also:

- promote the understanding of science by the public, policy makers and elected officials, and assist in the understanding of where science can benefit New Zealand
- undertake activities that enhance the use of science and evidence in policy formation and evaluation across government (DPMC (NZ) 2019, p. 1).

Chief Science Advisors are also housed in the major departments of New Zealand. Their role includes acting as in-house peer reviewers of evidence and conduits between their department and the science community (Jeffares et al. 2019, p. 63). The PMCSA has also recently established a *Chief Science Advisor Forum* (consisting of the PMCSA, Chief Science Advisors across government, and co-opted members). The remit of this Forum appears to be in the process of being established, but its draft terms of reference points to one of its functions being to ‘advance the use of science to benefit New Zealand through promoting the use of evidence to inform policy development, practice and evaluation’ (Office of the Prime Minister’s Chief Science Advisor (NZ) 2018, p. 1).

Requiring management responses to evaluation recommendations

Another mechanism to promote evaluation use in the policymaking process is to encourage or require managerial responses to evaluation recommendations, including whether recommendations are accepted, why or why not, and actions to be taken in response (OECD 2020b, p. 137).

The OECD found that the use of formal management responses and follow-up procedures for evaluation is relatively uncommon at a departmental level (OECD 2020b, p. 137). One example where such a process exists is in **Canada**, where evaluation reports must be

accompanied by a management response and action plan.⁸² In some countries, it is typical for evaluations to be responded to out of convention. For example, in the **United Kingdom**, it is common for ministers or managers to respond to evaluation results even though there is no formal requirement to do so (OECD 2020b, p. 139).

Whole-of-government responses to evaluations also exist in some countries. A notable example is **Japan**, where the government provides an annual report to the Diet (Japan's bicameral legislature) on policy evaluation, including how results have been reflected in policy planning and development (OECD 2020b, p. 137).

D.4 Approaches to evaluating policies and programs affecting indigenous peoples

This section explores two key questions:

- How do countries approach evaluating policies and programs targeted specifically at indigenous peoples?
- How do countries approach evaluating impacts and outcomes for indigenous peoples when evaluating mainstream policies and programs?

In contrast to previous sections (which have taken a cross-country approach to examining how countries approach evaluation generally), this section examines three countries — Canada, the United States and New Zealand. The Commission has focused on these countries in part because they are all colonial-settler nations with significant indigenous populations and governance and policymaking structures that are — relatively speaking — similar to Australia. A number of project participants also suggested that these countries could harbour useful lessons for the Commission as it designed the Indigenous Evaluation Strategy.⁸³

This is not to say, however, that other countries (with primary languages other than English) do not have other approaches to evaluating policies and programs affecting indigenous peoples. For example, in India, one of the functions of Tribal Research Institutes is to conduct research and evaluation studies⁸⁴ (Ministry of Tribal Affairs (India) 2020). And in Mexico, the legal framework governing the government indigenous authority, the National Institute of Indigenous Peoples (Instituto Nacional de los Pueblos Indígenas), requires them to work jointly with the National Council of Social Development Policy Evaluation (Consejo

⁸² This requirement is spelled out as part of the *Standard on Evaluation* (section C.2.2.6.9) of the Directive on Results (Government of Canada 2016a).

⁸³ These participants included: the Lowitja Institute (sub. 50); the National Aboriginal Community Controlled Health Organisation (sub. 95); and the Queensland Indigenous Family Violence Legal Service (sub. 25).

⁸⁴ Tribal Research Institutes are a network of 25 institutes administered by Indian state governments. In addition to research and evaluation, other functions of Tribal Research Institutes include: undertaking capability building and training for people and institutions working in tribal development; collecting baseline data; organising tribal festivals; conducting seminars and workshops; and providing planning inputs to state governments.

Nacional de Evaluación de la Política de Desarrollo Social, CONEVAL) to evaluate public policies, and the implementation of plans, programs, projects and government actions, as well as issue recommendations to guarantee full access to social rights. (CONEVAL 2019, p. 36) CONEVAL is:

... a decentralized public agency that coordinates the evaluation of the National Social Development Policy as well as other policies, programs, or interventions related to social development. CONEVAL provides guidelines for evaluations that must be followed by other public agencies that implement social programs. At the same time, CONEVAL can directly contract evaluations on social policies and programs, and provide recommendations for improvement based on available evaluations and research. (3ie International Initiative for Impact Evaluation nd)

CONEVAL also undertakes thematic evaluation reports that provide a general assessment of the performance of programs across a number of key social policy themes, including programs that impact indigenous people (for example, CONEVAL 2014).

Evaluation of policies and programs affecting indigenous peoples in Canada

The indigenous peoples of Canada comprise of the First Nations, Inuit and Métis. The relationship between First Nations, Inuit and Métis and the Government of Canada is governed by a number of treaties, laws (including the Constitution) and court decisions (Department of Justice Canada 2018; OECD 2020d). Unless there is negotiated self-government in place, most First Nations are currently governed by the Federal Government *Indian Act 1985* (Canada) (which establishes a limited form of local administration), although self-government is increasing. There are currently 25 self-government agreements (involving 43 communities) and about 50 current ‘negotiation tables’ across Canada. There are also some additional agreements relating to education (Crown-Indigenous Relations and Northern Affairs Canada 2019b).

Departmental responsibility for policies and programs affecting indigenous peoples in Canada is undergoing significant change. Until recently, Indigenous and Northern Affairs Canada was the federal department primarily responsible for meeting Canada’s commitments and obligations to indigenous peoples. However, in 2017, this department was dissolved and replaced with two new departments:

- *Indigenous Services Canada*, which works with partners to improve access to high-quality services for First Nations, Inuit and Métis
- *Crown-Indigenous Relations and Northern Affairs Canada*, which seeks to renew the relationship between Canada and First Nations, Inuit and Métis peoples and, amongst other things, oversees matters relating to treaties, agreements and self-governance (Government of Canada 2020c, 2020d, 2020a).

The main instrument governing the evaluation of policies and programs affecting indigenous Canadians (at a federal level) is the Policy on Results that applies to Canadian Government

agencies (section D.1). This means that both Indigenous Services Canada and Crown-Indigenous Relations and Northern Affairs Canada are required to identify a head of evaluation, prepare five year evaluation plans, evaluate their programs at least once every five years (consistent with the *Financial Administration Act 1985* (Canada)) and publish evaluation reports, including management responses, on their websites (as prescribed by the Treasury Board Secretariat).

The number of evaluations published by these departments is small (which most likely reflects that they been established relatively recently). At the time of writing this appendix, Indigenous Services Canada had published six evaluations (two of which are discussed in box D.6) (Indigenous Services Canada 2020a), while the Commission could not find any evaluations published by Crown-Indigenous Relations and Northern Affairs Canada. However, the predecessor of these agencies — Indigenous and Northern Affairs Canada — published over 100 evaluations between 2007 and 2017 (Crown-Indigenous Relations and Northern Affairs Canada 2019a), meaning the stock of evaluations of programs for indigenous peoples in Canada is considerable (at least at a federal level).

Some studies have used the large stock of evaluations to undertake meta-analyses to make observations about the quality of evaluations of indigenous policies and programs in Canada. Although these studies are somewhat dated (and therefore may not be representative of current practice), they are worth examining as they can yield broad insights into how effective evaluation in Canada has been in the past. For example:

- a 2013 study found that, while it was apparent that cultural sensitivity was gradually being integrated into Aboriginal⁸⁵ program evaluation in Canada, the integration of participatory approaches was ‘in sum, relatively limited’ and that ‘local populations hold little or no decision-making power over the evaluation process’ (Jacob, Desautels and University 2013, p. 23)
- a different study by the same authors in 2014 — this time focusing on the quality of evaluation reports — found that the evaluation of Aboriginal programs ‘is of good, or even excellent, quality’ (Jacob and Desautels 2014, p. 78) and that the then central Canadian evaluation policy (the predecessor to the current Policy on Results) had a definitive impact on the quality of evaluations (p. 62).

⁸⁵ The term ‘Aboriginal Canadians’ has sometimes been used to refer to the indigenous peoples of Canada (Anaya 2014). In Library and Archives Canada’s *Terminology Guide: Research on Aboriginal Heritage*, Aboriginal Peoples is defined as ‘a collective name for the original peoples of North America and their descendants’ (Library and Archives Canada 2012, p. 6). The studies cited have used the term ‘Aboriginal programs’.

Box D.6 Overview of two evaluations undertaken by Indigenous Services Canada

Evaluation of the First Nations and Inuit Home and Community Care Program

The First Nations and Inuit Home and Community Care Program (FNIHCC) involves contribution agreements with First Nation and Inuit communities (along with territorial governments) to fund the administration of nurses and personal care workers in over 450 First Nation and Inuit communities. An evaluation of the program was conducted in 2019, drawing on evidence from a variety of sources, including administrative data, a document and file review, interviews, and a web-based survey of (primarily) health directors, care coordinators and nurses.

The evaluation found that the program:

... has demonstrated considerable success over the past five years, with continued improvements largely due to effective program management and dedication of First Nation health directors and/or Home and Community Care coordinators and nurses. (2019, p. iii)

It also found that the priorities of the Government and indigenous communities 'generally aligned with respect to care provided under the FNIHCC' (2019, p. 5) but that policy and legislative gaps risked creating or maintaining disparities between indigenous and non-indigenous Canadians. It also found the extent to which the program met demand could not be measured because of constraints in data collected on demand for services.

The evaluation made four recommendations to improve the program and the evaluation report included a management response and action plan for these recommendations. Also included in the evaluation report was a discussion about evaluation questions and methodological limitations. A two-page *Results at a Glance* summary for the evaluation was also produced.

Evaluation of the On-Reserve Income Assistance Program

The On-Reserve Income Assistance Program is part of Canada's social safety net. It is a program of last resort, and aims to ensure that eligible individuals and families that live on-reserve have funds to cover their basic expenses of daily living, and to support their transition to self-sufficiency. An evaluation of the program was conducted in 2018. Evidence to inform the evaluation included interviews (including with chiefs and community leadership, service delivery staff, representatives of tribal councils and income assistance recipients), document and literature reviews, analysis of quantitative data and a financial review.

The evaluation found that the program addresses a continued need and is highly relevant to government objectives. But it found that the program's design constrained its ability to deliver on its desired outcomes:

The desired outcomes of [the] Income Assistance Program are limited by three core assumptions applied to its design: 1) that alignment with the provinces and territory of Yukon is appropriate; 2) that Income Assistance can promote attachment to the workforce without corresponding investment in active measures; and 3) that the program can operate successfully without meaningful and robust engagement with First Nations stakeholders. The evaluation finds evidence that these assumptions require reconsideration. (2018, p. 33)

The evaluation provided five recommendations, including that Indigenous Services Canada co-develop a new Income Assistance policy with First Nations groups (with other federal departments as collaborators). In their response to the evaluation, management concurred.

Sources: Indigenous Services Canada (2018, 2019, 2020b).

When ‘mainstream’ programs — that is, programs that provide services to all Canadians that are not necessarily targeted at First Nations, Inuit or Métis peoples — are evaluated, there appears to be no centrally determined requirement to explicitly consider the program’s impact on (or effectiveness for) indigenous Canadians. While there are examples of evaluations of mainstream programs that do provide some level of reporting of the program’s impacts on indigenous Canadians (for example, a 2019 evaluation on apprenticeship grants by Employment and Social Development Canada (2019)), this appears to be done on an ad-hoc basis.

Evaluation of policies and programs affecting indigenous people in Canada also occurs at a provincial and territorial level. However, there is evidence to suggest that the scale of evaluation is not as large as what occurs at the federal level. Jacob and Desautels (2014, p. 67) found that ‘by far the biggest player’ in evaluating indigenous programs in Canada was the then federal department (Indigenous and Northern Affairs Canada).

Evaluation of policies and programs affecting indigenous peoples in the United States

The indigenous peoples of the United States:

... include a vast array of distinct groups that fall under the generally accepted designation of Native Americans, which include American Indians and Alaska Natives; also included are the people indigenous to Hawaii, or Native Hawaiians. (Anaya 2012, p. 5)

Numerous instruments underpin relationships between Native American Tribes and the Government of the United States. On this relationship, the Tribal Epidemiological Centers (2013) stated:

Federally-recognized tribes of American Indians and Alaska Natives have a unique historic and legal government-to-government relationship with the U.S. government. This relationship has been given substance through numerous Supreme Court decisions, treaties, legislation, and Executive Orders, and entitles AI/ANs and their descendants to special rights, including health benefits. Under this trust responsibility, the U.S. has a legal obligation to protect Tribes’ assets and provide needed services to Indian people. (p. 9)

While several different departments run programs for Native Americans, two of the bigger service providers are the Department of the Interior (DOI) and the Department of Health and Human Services (HHS). The former houses the Bureau of Indian Affairs, while the latter houses the Indian Health Service (IHS) and the Administration for Native Americans (which is part of the Administration for Children and Families).

There appears to be no systematic, whole-of-government approach to evaluating policies and programs affecting Native Americans in the United States, and different departments appear to have different approaches.

The Administration for Native Americans, which — among other functions — provides grants to support locally determined projects designed to reduce or eliminate community

problems and achieve community goals — maintains a division of program evaluation and planning. It meets with about one-third of its grantees each year for the purposes of evaluating the impacts of their grants and to learn about the successes and challenges of grant recipients (ANA 2019a, 2019b, nd).

Evaluation of projects is informed by standard tools, and results are used to inform a report to Congress on the impact and effectiveness of the Administration's projects (ACF and ANA 2019; ANA 2019c). The parent agency of the Administration for Native Americans — the Administration for Children and Families (ACF) — also maintains an Office of Planning, Research and Evaluation. Its functions include evaluating existing programs, which can include examining impacts of ACF programs on or for Native Americans (with one example being an evaluation of a grant program to train individuals in health-related professions (Meit et al. 2016; OPRE 2020)). The Office also funds the Tribal Evaluation Institute, which works collaboratively with grant recipients under the federally-funded Tribal Home Visiting Program to evaluate how home visiting can improve outcomes in Tribal communities (Tribal Evaluation Institute 2016).

The IHS also maintains a planning, evaluation and research branch. It has developed an evaluation policy that outlines some guidance on how the evaluation of IHS projects should be undertaken. Amongst other things, the policy states that all new projects should include evaluation as part of program planning and development. This includes using five to ten per cent of program funds to support evaluation, and developing a theory of change for the program that outlines expected outcomes in the short, medium and long term (IHS nd).

The IHS also funds (with other agencies) Tribal Epidemiology Centers (TECs). TECs work in partnership with Tribes to improve the health and wellbeing of people in their Tribal community. There are 12 TECs, each covering a different geographic area (one TEC, based in Seattle, has a nationwide focus on urban Native Americans) (CDC nd). TECs have several functions including: data collection relating to, and progress monitoring of, health objectives; assisting Indian Tribes, tribal organisations and urban Indian organisations to identify highest-priority health objectives and services to meet these objectives; and making recommendations on how services could be targeted or improved. TECs also evaluate delivery, data and other systems that impact on the improvement of Indian health (Tribal Epidemiology Centers 2020). TECs have a lot of flexibility in how they go about fulfilling their duties. Their functions are to be provided '[i]n consultation with and on the request of Indian tribes, tribal organizations, and urban Indian organizations' and each TEC is unique in the types and range of activities they conduct and services they provide (CDC nd, p. 1).

The introduction and implementation of the *Foundations for Evidence-based Policymaking Act 2018* (US) (discussed above) may change the frequency or approach with which policies and programs affecting Native Americans are evaluated. Both the DOI and the HHS — as federal government agencies — appear to be subject to the requirements set out in the Act, including developing evaluation plans and appointing evaluation officers.

Evaluation of policies and programs affecting indigenous peoples in Aotearoa New Zealand

The Māori are the indigenous people of Aotearoa New Zealand. They are a Polynesian people, who arrived in Aotearoa New Zealand in the late 13th Century, over 300 years before the arrival of Europeans. In 1840, as British efforts to colonise New Zealand began, the British Crown and Māori chiefs signed the Treaty of Waitangi (Wilson 2020). Today the Treaty is widely accepted as a constitutional document that establishes and guides the relationship between Māori and the New Zealand Government, although its status in New Zealand law has been described as ‘less than settled’ (Ministry of Justice (NZ) 2020).

There appears to be no formal, publicly available central framework that New Zealand government agencies must apply when evaluating policies or programs affecting Māori. A number of agencies have produced guidelines for evaluation that involves or affects Māori, although many of these are over a decade old and the extent that they are still drawn upon is not clear (box D.7).

Box D.7 Guidelines for evaluation involving Māori

In 1999, Te Puni Kōkiri (the Ministry for Māori Development) published evaluation guidelines for New Zealand government agencies undertaking evaluations to collect quality information about Māori. The guidelines explore why evaluating for Māori is important and ethical issues that should be considered when undertaking evaluation involving Māori. The guidelines also step through the critical stages of an evaluation, including evaluation planning, design, analysis, reporting and communicating results. For each stage, the guidelines outline critical success factors and some commentary on what good practice looks like, common gaps and a checklist for evaluators to follow when completing each step.

In 2004, the Centre for Social Research and Evaluation (CSRE, which was part of the New Zealand Ministry of Social Development) published *Guidelines for Research and Evaluation with Māori* to assist staff from the CSRE and Ministry researchers to undertake projects that require input from Māori. The guidelines are structured around six practice principles: planning for Māori involvement; engaging with Māori participants and stakeholders; developing effective and appropriate methodologies; protecting knowledge; encouraging reciprocity; and supporting Māori development. A set of actions for evaluators is provided for each principle, as well as supporting advice and commentary to help staff operationalise these principles and to provide more information. The guidelines also outline a set of core competencies and expected conduct for research contractors.

In 2008, the Social Policy Evaluation and Research Committee (SPEaR) published good practice guidelines ‘to enhance the standard of research and evaluation practice across the social sector as a whole’ (SPEaR 2008, p. 5). These guidelines articulate five principles that uphold good practice: respect; integrity; responsiveness; competency; and reciprocity. The guidelines also highlight the importance of utility and data sharing in research and evaluation. The guidelines then focus on the application of the principles in several contexts, with one being in research and evaluation involving Māori, including drawing on the experiences of those involved with research and evaluation to describe what good practice looks like.

Sources: CSRE (2004); SPEaR (2008); TPK (1999).

More recently, Te Arawhiti (the Office for Māori Crown Relations) — established in 2018 — produced a framework and guidelines on how agencies should engage with Māori, including as part of the policymaking process (Te Arawhiti (Office for Māori Crown Relations) nd). These documents provide guidance to agencies on: defining the kaupapa (the policy, purpose or matter on which engagement is sought); who to engage with; how to engage and when; and how to develop an engagement strategy.

Te Arawhiti has also developed a Māori Crown Relations Capability Framework. The Framework aims to position the public service to support the Māori Crown relationship, enable government to consistently meet its obligations under the Treaty of Waitangi, and achieve a uniquely New Zealand public service that is able to best serve all New Zealanders (Te Arawhiti (Office for Māori Crown Relations) 2019a, p. 2). The framework has several components — including an individual and organisational capability component — that set out different levels of capability maturity and provide guidance on steps that can be taken to improve capability (Te Arawhiti (Office for Māori Crown Relations) 2019c, 2019b).

Another recent development in Aotearoa New Zealand is a change to the monitoring role of Te Puni Kōkiri (TPK). Te Puni Kōkiri (or the Ministry of Māori Development) is a government agency with a range of roles and functions including: leading policy advice to government on strategic issues affecting Māori wellbeing and development; sustaining local relationships to inform policy and service design and connecting Māori with government agencies; and influencing and supporting public sector agencies and other organisations on what will make the biggest difference, based on evidence (TPK 2020a).

TPK also has a monitoring function (underpinned by legislation). This function has taken different forms over time (Minister for Māori Development 2019b), but was recently refreshed to cover three streams:

1. monitoring wellbeing outcomes, through a periodic report to identify elements of progress towards Māori wellbeing
2. monitoring progress of government priorities by tracking gains for Māori across government priorities
3. monitoring policy, program and service effectiveness by undertaking effectiveness reviews of priority policies, programs and services to understand their effectiveness for Māori and their contribution towards Māori wellbeing (TPK 2020b).

The Minister for Māori Development recently said that it was her expectation that TPK's monitoring function will:

... increase transparency of, and accountability for, the effectiveness of government agencies for Māori by monitoring the impact of their strategies, policies and services on Māori wellbeing outcomes. This will be a critical opportunity to inform policy and build the evidence base of what works to improve Māori wellbeing outcomes. (2019a, p. 6)

In late 2019, TPK completed a pilot monitoring review of the New Zealand Ministry of Social Development's employment services and programs in Auckland. Along with

assessing the contribution of the Ministry's services and programs to Māori wellbeing, the pilot provided an opportunity for TPK to trial its review process and assess the extent that it is fit-for-purpose and replicable for future reviews (TPK 2019, 2020b).

In addition to TPK's national and government-wide monitoring role, monitoring outcomes for Māori sometimes also occurs at more localised levels, such as in Auckland through the Independent Māori Statutory Board (box D.8).

Box D.8 **The role of the Independent Māori Statutory Board**

The Independent Māori Statutory Board has specific responsibilities and powers under the *Local Government (Auckland Council) Amendment Act 2010* (NZ). The Board provides direction and guidance to Auckland Council on issues affecting Māori. The Board also has appointees on many Council committees.

The Board has a statutory responsibility for promoting 'Issues of Significance' to Māori in Tāmaki Makaurau (Auckland) and to monitor the Council's performance in responding to these issues. It produces the Māori Plan, which outlines what Māori in the Tāmaki Makaurau region say is important to them and 'provides a framework for understanding Māori development aspirations and monitoring progress towards desired cultural, economic, environmental and social outcomes' (Independent Māori Statutory Board 2017, p. 4). The Board also measures and reports progress in Māori wellbeing.

Every three years, the Board conducts an audit to assess the Council's performance in acting in accordance with the Te Tiriti o Waitangi (the Treaty of Waitangi) and its responsibilities to Māori in Tāmaki Makaurau. The Board has also commissioned reports examining the effectiveness of the Council's systems for planning and expenditure on projects to improve Māori outcomes, and has developed business cases for where the Council should focus to deliver on its strategy for Māori.

Source: Independent Māori Statutory Board (2017, nda, ndb, ndc).

There is also a growing emergence and awareness of Māori evaluation theories, such as Kaupapa Māori evaluation (or KME) in New Zealand. KME, and Māori evaluation more broadly, can be characterised as evaluation that is:

- controlled and owned by Māori
- conducted for the benefit of Māori (although it may 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), quoted in Moewaka Barnes (2013, p. 165)).

Kaupapa Māori originated 'out of concern for the unjust and harmful impacts endured by Māori at the hands of non-Māori researchers' (Carlson, Moewaka Barnes and McCreanor 2017, p. 70). Cram, Pipi and Paipa described KME as being about:

... the reclaiming of knowledge creation mechanisms. This can be done by growing Māori-centred explanations 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)

Carlson, Moewaka Barnes and McCreanor described Kaupapa Māori theory, research, action and evaluation as:

... critically oriented, methodologically eclectic, and encourage rigour while celebrating diversity, community-centred approaches and the expanding sense of understanding of the realms of te ao Māori (the Māori world). (2017, p. 70)

Drawing on their experiences undertaking KME evaluations, Cram, Pipi and Paipa (2018) stated that in terms of methods, they ‘are seeking ones that are right for answering evaluation questions about making a positive difference for Māori’ (p. 69). But regardless of method, engagement processes with stakeholders are often similar, beginning with whakawhanaungatanga, or the process to establish connections between evaluators and stakeholders. Cram, Pipi and Paipa also articulate a community-up approach to evaluator conduct, consisting of seven values to guide engagement with communities and organisations when undertaking evaluation (box D.9).

Box D.9 A ‘community-up’ approach to defining evaluation conduct

Cram, Pipi and Paipa describe a community-up approach to evaluation conduct that has seven values, each with their own guideline. These are:

- Aroha ki te tangata — Respect people — allow them to define their own space and meet on their terms
- He kanohi kitea — Meet people face-to-face, and also be a face that is known to and seen within a community
- Titiro, whakarongo ... kōrero — Look and listen (then maybe speak) — develop an understanding in order to find a place from which to speak
- Manaaki ki te tangata — Share, host and be generous
- Kia tūpato — Be cautious — be politically astute, culturally safe, and reflective about insider/outsider status
- Kaua e takahia te mana o te tangata — do not trample on the ‘mana’ or dignity of a person
- Kia māhaki — Be humble — do not flaunt your knowledge; find ways of sharing it.

Source: Cram, Pipi and Paipa (2018, pp. 70–72).

While the application of KME can involve compromise by evaluators (funders of KME can determine evaluation parameters that do not necessarily align with a localised focus for KME) (Carlson, Moewaka Barnes and McCreanor 2017, p. 74), there are several recent evaluations that identify as applying KME. These include evaluations into programs centred on family safety (Wehipeihana 2019b) and preventing accidents in the home (Hayward and Villa 2015).

Māori evaluators have also formed a network called Mā te Rae (the Māori Evaluation Association). Founded in 2015, Mā te Rae was established by Māori for Māori ‘to advance the social, cultural and economic development of Iwi Māori through participation in and contribution to quality evaluation’ and seeks to ‘mobilise evaluation as a tool for transformation for Iwi Māori’ (Mā te Rae 2015). Mā te Rae provides a space for Māori evaluators to connect in their own way and plays an important role in supporting and building capacity for Māori-led evaluation (Mā te Rae pers. comm., 8 May 2020).

One of the recent activities of Mā te Rae was to host an Indigenous Peoples’ Conference on Evaluation, which explored several themes, including how to use traditional knowledge in evaluation, and ‘claiming the space’ (self-determination) (Mā te Rae nd). Mā te Rae also maintains connections to international groups, including the EVALINDIGENOUS network — a multi-stakeholder partnership to ‘advance the contribution of Indigenous evaluation to global evaluation practice’ by recognising different world views and valuing the strengths of indigenous evaluation practices (EvalPartners 2020).

References

- 3ie International Initiative for Impact Evaluation nd, *National Council for the Evaluation of Social Development Policy (CONEVAL)*, <https://www.3ieimpact.org/about-us/members/national-council-evaluation-social-development-policy-coneval> (accessed 30 September 2020).
- Aboriginal Affairs NSW 2018, *OCHRE: Five Years On*, Aboriginal Affairs NSW, Sydney.
- Aboriginal Affairs (NSW) and NSW Department of Education 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.
- Ackhurst, M., Polvere, R.-A. and Windley, G. 2017, *Indigenous participation in VET: understanding the research*, National Centre for Vocational Education Research.
- 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, *Code of Ethics*, Melbourne.

-
- 2013b, *Evaluators' Professional Learning Competency Framework*, Melbourne.
- 2013c, *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*.
- AIATSIS (Australian Institute of Aboriginal and Torres Strait Islander Studies) 2012, *Guidelines for Ethical Research in Australian Indigenous Studies*, Canberra.
- 2020a, *A Guide to Applying the AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*, Canberra.
- 2020b, *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*, Canberra.
- 2020c, *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).
- 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).
- 2020, *Knowledge Circle*, Text, Australian Institute of Family Studies, <https://aifs.gov.au/past-projects/knowledge-circle> (accessed 8 October 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.
- 2010, *National Best Practice Guidelines for Collecting Indigenous Status in Health Data Sets*, IHW 29, Canberra.

-
- 2017, *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).
- 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.
- ALRC (Australian Law Reform Commission) 2015, *Connection to Country: Review of the Native Title Act 1993 (Cth) - Final Report*, Sydney.
- 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.
- Amies, M. 1995, ‘Program Evaluation: A Commonwealth Perspective. Where are we now?’, in Guthrie, J. (ed), *Making the Australian Public Sector Count in the 1990s*, IIR Conferences Pty Ltd, Sydney.
- 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.
- 2014a, *Initiatives to Support the Delivery of Services to Indigenous Australians*, Performance Audit, No. 45 2013-14, Canberra.
- 2014b, *Public Sector Governance*, 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, No. 47 2018-19, Canberra.
- 2019c, *Evaluating Aboriginal and Torres Strait Islander Programs*, Auditor-General Report No. 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).
- 2020c, *APS Data Professional Stream*, <https://www.apsc.gov.au/aps-data-professional-stream> (accessed 30 September 2020).
- 2020d, *APS Digital Professional Stream*, <https://www.apsc.gov.au/aps-digital-professional-stream> (accessed 30 September 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 Indigenous HealthInfoNet 2020, *Our Work — What We Do*, Edith Cowan University, <https://healthinfonet.ecu.edu.au/about/our-work/what-we-do/> (accessed 16 April 2020).
- Australian Institute of Criminology and Council of Attorneys-General 2016, *Indigenous Justice Clearinghouse*, <https://www.indigenousjustice.gov.au/> (accessed 25 August 2020).
- 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.
- Barnes, H.M. 2000, ‘Kaupapa Māori: explaining the ordinary’, *Pacific Health Dialogue*, vol. 7, no. 2, pp. 13–16.
- Bazalgette, L. 2020, *A Practical Guide for Establishing an Evidence Centre*, Alliance for Useful Evidence, Nesta, London.
- 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.
- BetterEvaluation 2010, *Stories*, BetterEvaluation, <https://www.betterevaluation.org/en/evaluation-options/stories> (accessed 8 October 2020).
- 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.
- Blackwell, K. 2019, *Indigenous languages and government service delivery: what do evaluations tell us? (Unpublished research paper)*, 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.

-
- 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 Ifrom Other Influences (Supplementary Magenta Book Guidance)*, December, United Kingdom.
- Cargo, M., Potaka-Osborne, G., Cvitanovic, L., Warner, L., Clarke, S., Judd, J., Chakraborty, A. and Boulton, A. 2019, ‘Strategies to support culturally safe health and wellbeing evaluations in Indigenous settings in Australia and New Zealand: a concept mapping study’, *International Journal for Equity in Health*, vol. 18, no. 1, p. 194.
- Carlson, T., Moewaka Barnes, H. and McCreanor, T. 2017, *Kaupapa Māori Evaluation: A Collaborative Journey*, New Zealand Council for Educational Research.
- CBPATSISP (Centre of Best Practice in Aboriginal and Torres Strait Islander Suicide PreventionCentre of Best Practice in Aboriginal and Torres Strait Islander Suicide Prevention) 2020, *Clearing House*, Centre of Best Practice in Aboriginal and Torres Strait Islander Suicide Prevention, <https://www.cbpatsisp.com.au/clearing-house/> (accessed 8 October 2020).

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.

— and Centre for Aboriginal Health 2019, *Mid-Term Evaluation of the NSW Aboriginal Health Plan 2013-2023*, 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 2020, *Priority Reforms*, <https://www.closingthegap.gov.au/priority-reforms> (accessed 21 August 2020).

COAG (Council of Australian Governments) 2009, *National Indigenous Reform Agreement (Closing the Gap)*.

— 2016, *Prison to Work*, Canberra.

— 2018, *COAG Statement on the Closing the Gap Refresh 12 December 2018*.

— 2020, *National Legal Assistance Partnership 2020–2025*, Canberra, <https://www.ag.gov.au/system/files/2020-06/National-Legal-Assistance-Partnership-Agreement.pdf> (accessed 1 September 2020).

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.

-
- Coalition of Peaks 2020a, *Media Release, Coalition of Peaks, 30 July 2020*, <https://coalitionofpeaks.org.au/signing-of-new-national-agreement-on-closing-the-gap-marks-historic-turning-point-in-relations-between-aboriginal-and-torres-strait-islander-people-and-governments/> (accessed 27 August 2020).
- 2020b, *Welcome! We are the Coalition of Peaks*, <https://coalitionofpeaks.org.au/> (accessed 3 October 2020).
- Cobb-Clark, D.A. 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).
- Commonwealth of Australia 2019, *Budget Measures — Budget Paper No. 2, Budget 2019-20*, Canberra.
- CONEVAL (Consejo Nacional de la Evaluación de la Política de Desarrollo Social) 2014, *Evaluación Integral del Desempeño de los Programas Federales y Acciones Dirigidas a Población Indígena 2012-13*, Mexico.
- 2019, *Calidad de las evaluaciones del CONEVAL: elementos institucionales y methodológicos*, Mexico.
- 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.
- Cornish, L. 2020, *Independent evaluation lost as DFAT ‘adjusts’ its structure*, Devex, <https://www.devex.com/news/independent-evaluation-lost-as-dfat-adjusts-its-structure-98102> (accessed 25 September 2020).
- 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.
- Cox Inall Ridgeway 2019, *Review of the Indigenous Legal Assistance Program (ILAP) 2015–2020*, Final report.
- 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-circnac.gc.ca/eng/1100100032275/1529354547314#chp1> (accessed 30 April 2020).
- CSRE (Centre for Social Research and Evaluation (Te Pokapū Rangahaū Arotake Haporī)) 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.
- Dawkins, Z., Brookes, M., Middlin, K. and Crossley, P. 2011, *County Koori Court: Final Evaluation Report*, County Court of Victoria and the Victorian Department of Justice, Melbourne.
- 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 for Health and Ageing, Government of South Australia 2014, *Aboriginal Health Impact Statement Policy Guideline*, Adelaide.
- 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 and Training (Vic) 2018, *Research and Evaluation Register*, <https://www.education.vic.gov.au/about/research/Pages/researchreg.aspx> (accessed 25 August 2020).
- 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 1993, *Annual Report 1992-93*, Canberra.
- 1994a, *Doing Evaluations: A practical guide*, Canberra.
- 1994b, *The Use of Evaluation in the 1994-95 Budget*, Canberra.
- 1995, *Register of Published Evaluation Reports; Number 8*, September, Canberra.
- 2017, *Annual Performance Statements for Commonwealth Entities*, Resource Management Guide No. 134.

-
- 2019, *Commonwealth Procurement Rules*, Canberra.
- Department of Health, Government of Western Australia 2017, *Aboriginal Health Impact Statement and Declaration for Department of Health and Health Service Providers*, Perth.
- 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 (NT) 2020, *Northern Territory Government Program Evaluation Framework*, May, Darwin.
- Department of Treasury (WA) 2015, *Evaluation Guide*, Program Evaluation Unit, Perth.
- 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, *Evaluation in DFAT: Improving Quality and Use*, Canberra.
- 2020, *Aid evaluation plans*, <https://www.dfat.gov.au/development/performance-assessment/aid-evaluation/aid-evaluation-plans> (accessed 1 October 2020).
- 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) 2017, *Evaluation Strategy 2017–2021*, Australian Government, Canberra.
- nd, *DIIS Evaluation Plan 2018-22*, Canberra.
- Dillon, M.C. 1992, ‘Program evaluation and monitoring in Aboriginal and Torres Strait Islander affairs: A strategic overview’, in O’Fairchellaigh, C. and Ryan, B. (eds), *Program Evaluation and Performance Monitoring: an Australian Perspective*, Macmillan, South Melbourne, pp. 91–103.
- DIRD (Department of Infrastructure and Regional Development) 2016a, *Evaluation Resource Booklet 2016–2021*, Canberra.
- 2016b, *Evaluation Strategy 2016–2021*, 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).
- DITRDC (Department of Infrastructure, Transport, Regional Development and Communications) 2020, *Regional Australia Impact Statements Guidance*, Canberra.

-
- 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.
- 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, Canberra.
- DoH (Department of Health) 2016, *Evaluation Strategy 2019-2019*.
- 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*.
- DPC (Vic) (Department of Premier and Cabinet (Vic)) 2019, *Aboriginal-led accountability of government*, <https://engage.vic.gov.au/aboriginal-led-accountability> (accessed 29 September 2020).
- DPC (WA) (Department of the Premier and Cabinet (WA)) 2018, *An Office for Advocacy and Accountability in Aboriginal Affairs in Western Australia Discussion Paper*, Discussion Paper, WA Government, Perth.
- 2020, *Aboriginal Advisory Council of Western Australia*, Western Australian Government, <https://www.wa.gov.au/organisation/department-of-the-premier-and-cabinet/aboriginal-advisory-council-of-western-australia> (accessed 8 October 2020).
- DPMC (Department of the Prime Minister and Cabinet) 2017, *2017-18 Budget: Strengthening Indigenous Research and Evaluation*, 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, *An evaluation of the first two years of the Community Development Programme — Summary*, 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.
- 2020a, *Australian Government RIS Preliminary Assessment Form: Is a RIS required?*, Canberra.
- 2020b, *Data Availability and Transparency Bill 2020 Exposure Draft Consultation Paper*, Canberra.
- DPMC (NZ) (Department of the Prime Minister and Cabinet (NZ)) 2019, *Terms of Reference — Prime Minister's Science Advisory Committee*.
- Drawson, A.S., Toombs, E. and Mushquash, C.J. 2017, 'Indigenous Research Methods: A Systematic Review', *International Indigenous Policy Journal*, vol. 8, no. 2, <https://ojs.lib.uwo.ca/index.php/iipj/article/view/7515> (accessed 11 October 2020).
- Dreise, T. and Mazurski, E. 2018, *Weaving Knowledges: 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.
- 2019, *Data Exchange Policy*, <https://dex.dss.gov.au/policy> (accessed 12 May 2020).
- DTF (VIC) (Department of Treasury and Finance (Victoria)) 2013, *Economic Evaluation for Business Cases Technical Guidelines*, Victorian Government, Melbourne.
- 2019, *The Resource Management Framework for Victorian Government Departments*, Government of Victoria, Melbourne.
- Dudgeon, P., Bray, A., Darlaston-Jones, D. and Walker, R. 2020, *Aboriginal Participatory Action Research: an Indigenous Research Methodology Strengthening Decolonisation and Social and Emotional Wellbeing*, Melbourne.
- 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) 2019, *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.
- 2020a, *Memorandum for Heads of Executive Departments and Agencies - Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices*, Washington D.C.
- 2020b, *Preparation, Submission, and Execution of the Budget*, July, 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.
- Fisher, R. and Altman, J. 2020, *The world’s best fire management system is in northern Australia, and it’s led by Indigenous land managers*, The Conversation, <https://theconversation.com/the-worlds-best-fire-management-system-is-in-northern-australia-and-its-led-by-indigenous-land-managers-133071> (accessed 6 October 2020).
- 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.
- Fox, S. and Emerson, L. 2015, ‘Paper 2 - Effective interventions to promote a healthy start in life: Evidence-based programs for Aboriginal and Torres Strait Islander children and their families (Summary)’, *Good Beginnings: Getting it right in the early years*, The Lowitja Institute, Melbourne.
- Frydenberg, J. 2018, *First full-time Indigenous Policy Evaluation Commissioner of the Productivity Commission*, <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. 2020a, *Open Letter to the Australian Public Service – Friday 4 September 2020*, Australian Government, Canberra.

-
- and — 2020b, *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.
- 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.csps-efpc.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).
- nd, *About us*, <https://www.gov.uk/government/organisations/civil-service/about> (accessed 30 September 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.
- 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.
- 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*, <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).
- Higgins, D. and Davis, K. 2014, *Law and Justice: Prevention and Early Intervention Programs for Indigenous Youth*, Resource sheet no. 34, Closing the Gap Clearinghouse.

-
- Hill, R., Pert, P.L., Davies, J., Robinson, C.J., Walsh, F. and Falco-Mammone, F. 2013, *Indigenous Land Management in Australia: Extent, scope, diversity, barriers and success factors*, CSIRO Ecosystem Sciences, Cairns.
- HM Treasury nda, *Government Analytical Evaluation Capabilities Framework*, London.
- ndb, *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*, Canberra.
- 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.
- iap2 (International Association for Public Participation) 2018, *IAP2 Spectrum of Public Participation*, https://iap2.org.au/wp-content/uploads/2020/01/2018_IAP2_Spectrum.pdf (accessed 27 April 2020).
- IBA (Indigenous Business Australia) 2020, *Impact Framework: Understanding the Economic, Social and Cultural Impacts of IBA's Work*, 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, *Publications*, <https://www.imsb.maori.nz/publications/auckland-council-performance/> (accessed 20 September 2020b).
- ndc, *What we do*, <https://www.imsb.maori.nz/what-we-do/issues-of-significance/> (accessed 21 April 2020c).
- Indigenous Services Canada 2018, *Evaluation of the On-Reserve Income Assistance Program*, Evaluation, Performance Measurement and Review Branch, Canada.
- 2019, *Evaluation of the First Nations and Inuit Home and Community Care Program*, Evaluation, Performance Measurement and Review Branch, Canada.
- 2020a, *Evaluation reports*, <https://www.sac-isc.gc.ca/eng/1550775205150/1550775238874> (accessed 28 April 2020).
- 2020b, *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 University, L. 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) 2019, *Partnership Agreement on Closing the Gap 2019-2029*, Canberra.
- 2020, *National Agreement on Closing the Gap*, 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.
- Jowell, R. 2003, *Trying It Out, The Role of ‘Pilots’ in Policy-Making, Report of a Review of Government Pilots*, Government Chief Social Researcher’s Office, Cabinet Office, London.

-
- Katz, I., Newton, B.J., Bates, S. and Raven, M. 2016, *Evaluation Theories and Approaches; Relevance for Aboriginal Contexts*, Social Policy Research Centre, UNSW, Sydney.
- Kelaher, 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*, Maraguka and Just Reinvest NSW, Sydney.
- 2019, *Final Report for the Evaluation of the Individual Placement and Support Trial*, Department of Social Services, Canberra.
- LaFrance, J., Nichols, R. and Kirkhart, K.E. 2012, ‘Culture writes the script: On the centrality of context in indigenous evaluation’, *New Directions for Evaluation*, vol. 2012, no. 135, pp. 59–74.
- 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.
- Liamputpong, 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.
- Lowitja Institute 2020, *Knowledge translation*, <https://www.lowitja.org.au/page/research/knowledge-translation> (accessed 30 April 2020).
- Luetjens, J. and ‘t Hart, P. 2019, *Governing by looking back: learning from successes and failures*, Australia and New Zealand School of Government.

-
- 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, ECD Working Paper Series, 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.
- Maiam nayri Wingara Indigenous Data Sovereignty Collective and AIGI (Maiam nayri Wingara Indigenous Data Sovereignty Collective and Australian Indigenous Governance Institute) 2018a, *Indigenous Data Sovereignty: Communiqué*, Indigenous Data Sovereignty Summit, Canberra.
- 2018b, *Indigenous Data Sovereignty. Data for Governance: Governance of Data: Briefing Paper: 2018*.
- 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.
- 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.
- McGregor, D., Restoule, J. and Johnston, R. (eds) 2018, *Indigenous Research: Theories, Practices, and Relationships*, Canadian Scholars Pr.
- 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.
- 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 2019a, *Positioning Te Puni Kōkiri for Strategic Impact*.
- 2019b, *Te Puni Kōkiri's Monitoring Function: Improving State Sector Performance*.
- Ministry of Health, NSW Government 2017, *Aboriginal Health Impact Statement: Policy Directive*, Sydney.
- 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).
- Ministry of Tribal Affairs (India) 2020, *TRI*, <https://tribal.nic.in/tri.aspx#:~:text=The%20first%20such%20Institute%20was,other%20states%20were%20also%20covered>. (accessed 13 September 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. 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.
- Morley, S. 2014, *Success factors for Indigenous entrepreneurs and community-based enterprises*, Resource sheet no. 30, Closing the Gap Clearinghouse, Canberra.
- 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.
- 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.
- 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) 2019a, *2019-20 Annual Evaluation Work Plan*, Canberra.
- 2019b, *Indigenous Evaluation Committee Terms of Reference*, Australian Government, Canberra, <https://www.niaa.gov.au/sites/default/files/publications/iec-terms-reference.pdf> (accessed 8 March 2020).
- 2020a, *Closing the Gap*, <https://www.niaa.gov.au/indigenous-affairs/closing-gap> (accessed 26 August 2020).
- 2020b, *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.
- Nickerson, R. 1998, ‘Confirmation bias: a ubiquitous phenomenon in many guises’, *Review of General Psychology*, vol. 2, no. 2, pp. 175–220.
- Ninti One Limited 2017, *About CRC-REP*, Cooperative Research Centre — Remote Economic Participation, <https://old.crc-rep.com/about-us> (accessed 8 October 2020).
- 2020, *Remote Australia Online — Search*, Ninti One, <https://www.nintione.com.au/publications/remote-australia-online/search/> (accessed 8 October 2020).
- 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).

Nous Group and CAYLUS 2017, *Investing in the Future: The Impact of Youth Programs in Remote Central Australia: a Social Return on Investment (SROI) Analysis*.

NSW Government nd, *Aboriginal Health Impact Statement*, <https://www.aci.health.nsw.gov.au/about-aci/cultural-respect/aboriginal-health-impact-statement> (accessed 24 August 2020).

NSW Ombudsman 2019, *OCHRE Review Report*, Sydney.

NSW Treasury 2018, *Program Evaluation*, Treasury Circular, TC 18-03, NSW Government, Sydney.

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.

— 2020a, *Building Capacity for Evidence-Informed Policy-making: Lessons from Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris.

— 2020b, *Improving Governance with Policy Evaluation - Lessons from Country Experience*, OECD Public Governance Reviews, Paris.

— 2020c, *Institutionalisation, Quality and Use of Policy Evaluation: Lessons from Countries Experience*, Forthcoming, Paris.

— 2020d, *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).

— nda, *Learning what works (and what doesn't)*, <https://oes.gsa.gov/> (accessed 19 September 2020a).

— ndb, *Mission*, <https://oes.gsa.gov/about/> (accessed 27 April 2020b).

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.

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.
- 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: A Results and Delivery Culture in the Government of Canada’.
- 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.
- 2017, *Evaluation Flash Cards: Embedding Evaluative Thinking in Organizational Culture*, St. Paul.
- , McKegg, K. and Wehiipeihana, 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 (Productivity Commission) 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.
- 2016, *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.
- 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.
- Policy Profession (UK) 2019, *Policy Profession Standards - a Framework for Professional Development*, London.
- 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.
- 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.

RCIADIC (Royal Commission into Aboriginal Deaths in Custody) 1991, *Final Report*, 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.

Rogers, P. 2014, *Week 19: Ways of framing the difference between research and evaluation*, https://www.betterevaluation.org/en/blog/framing_the_difference_between_research_and_evaluation#:~:text=Research%20is%20about%20being%20empirical,interview%20data%20which%20collects%20descriptions. (accessed 9 January 2020).

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.

Schoenfeld, 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.

-
- Senior Officers Working Group 2015, *Investigation into Indigenous Land Administration and Use*, Report to the Council of Australian Governments, 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, C. and Fox, S. 2015, ‘Paper 3 - Effective interventions to promote a healthy start in life: Evidence-based programs for Aboriginal and Torres Strait Islander children and their families (Report)’, *Good Beginnings: Getting it right in the early years*, The Lowitja Institute, Melbourne.
- 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.
- SNAICC (Secretariat of National Aboriginal and Islander Child Care) 2012, *Opening Doors Through Partnerships: Practical Approaches to Developing Genuine Partnerships That Address Aboriginal and Torres Strait Islander Community Needs*, Victoria.
- 2020, *Creating Change Through Partnerships: An Introductory Guide to Partnerships Between Aboriginal and Torres Strait Islander and Non-Indigenous Organisations in Child and Family Services*.
- 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.
- Stern, E. 2015, *Impact Evaluation: A Guide for Commissioners and Managers*, Big Lottery Fund, Bond, Comic Relief and the UK Department for International Development, London.
- Stewart, J. 2014, *Developing a culture of evaluation and research*, Child Family Community Australia (CFCA) Paper, 28, Australian Institute of Family Studies, Melbourne.
- 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).
- TFMI (Task Force on Management Improvement) 1993, *The Australian Public Service Reformed: An Evaluation of a Decade of Management Report*, Canberra.
- 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.
- 2019, *Te Puni Kōkiri Monitoring Report: Review of Ministry of Social Development Employment Services and Programmes in Auckland*, Wellington.
- 2020a, *Roles and functions*, <https://www.tpk.govt.nz/en/mo-te-puni-kokiri/what-we-do/our-role-and-functions> (accessed 29 April 2020).
- 2020b, *State sector effectiveness*, <https://www.tpk.govt.nz/en/a-matou-kaupapa/state-sector-effectiveness> (accessed 20 September 2020).

-
- Treasury (QLD) 2020, *Queensland Government Program Evaluation Guidelines, Second Edition*, Brisbane.
- 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).
- Tribal Evaluation Institute 2016, *Evaluate impact*, <http://www.tribaleval.org/evaluate-impact/> (accessed 20 September 2020).
- 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*.
- UNEG (United Nations Evaluation Group) 2016a, *Evaluation Competencies for the UN System for Evaluators, Managers and Commissioners of Evaluation*, United Nations Evaluation Group, New York.
- 2016b, *UNEG Norms and Standards for Evaluation*, United Nations Evaluation Group, New York.
- United Nations Economic and Social Council 2019, *Traditional knowledge: generation, transmission and protection*, New York.
- University of Newcastle 2016, *48 Hour Follow Up Evaluation: Final Report*.
- WA Council of Social Service 2017, *Co-Design Toolkit*.
- Walker, R., Ballard, J. and Taylor, C. 2002, *Investigating appropriate evaluation methods and indicators for Indigenous housing programs*, Australian Housing and Urban Research Institute, Melbourne.
- 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.
- Walter, M. 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.

-
- 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.
- Wehiipeihana, 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.
- West, R., Stewart, L., Foster, K. and Usher, K. 2012, ‘Through a Critical Lens: Indigenist Research and the Dadirri Method’, *Qualitative Health Research*, <https://journals.sagepub.com/doi/10.1177/1049732312457596> (accessed 8 October 2020).
- 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. 2009, *Theory-based impact evaluation: Principles and Practice*, International Initiative for Impact Evaluation, New Delhi.
- , 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, 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 Encyclopaedia of New Zealand, <http://www.TeAra.govt.nz/en/history> (accessed 5 October 2020).

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 Louden, 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.