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[1] 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 2020a, p. 32).

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 2020a, p. 36).

---

[1] This appendix draws on a forthcoming cross-country study undertaken by the OECD entitled *Institutionalisation, Quality and Use of Evaluation: Governance: Lessons From Countries Experience* (OECD 2020a).

The frameworks that have been developed share a number of common elements. For example, of the frameworks developed by OECD countries:

- 88 per cent outlined the responsibilities of government institutions with respect to 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)
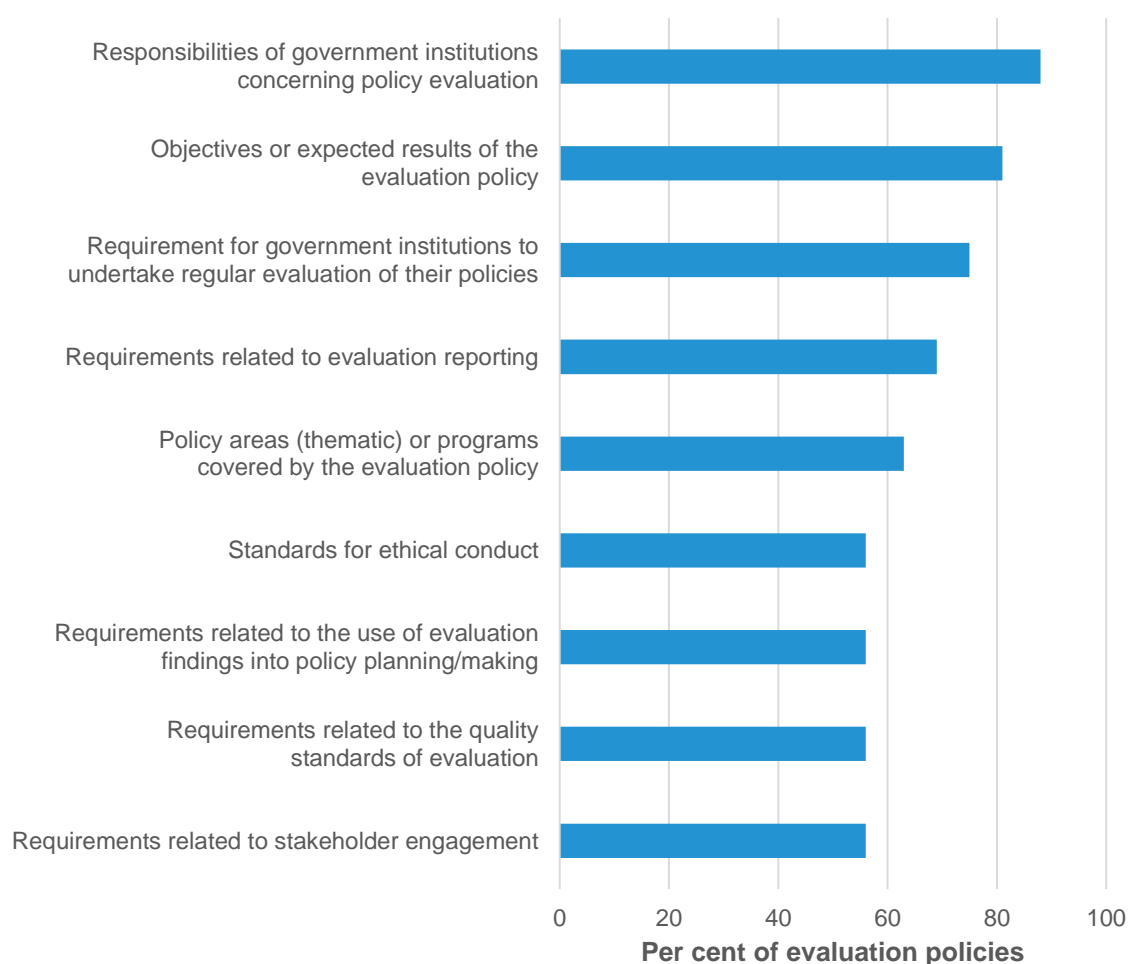
Data source: OECD (2020a, p. 37).

Canada's *Policy on Results* provides a useful example of an evaluation policy that embodies 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 with respect to 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 publicly 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 year[2] (Government of Canada 2016a, 2016b).

The Policy 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).

---

2   The requirement that ongoing programs be reviewed is also set out in s. 42.1 of the *Financial Administration Act 1985* (Canada).

Figure D.2   Roles required under the *Policy on Results*

	Deputy Department 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).

Implementation of — and compliance with — the Policy is overseen by the Canadian Treasury Board Secretariat.[3] 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

- 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).

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).

---

[3] The Treasury Board Secretariat is a central government agency. 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; TBS 2018).

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 _Standard_ 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, including by identifying 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).

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 evidence building activities in US federal agencies (Executive Office of the President (US) 2019b s. 290.2). 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) 2019b s. 290.6).

Under the Act, agencies are required to list (among other things):

- 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.[4]

The Act also requires agencies to *issue an evaluation plan* for each fiscal year that describes the evaluation activities it plans to conduct.[5] 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) 2019a, pp. 34–35).

The Act also obliges agency heads to designate a senior employee as the *agency's evaluation officer*.[6] 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.[7]

---

[4] These requirements are outlined in s. 312 of the *Foundations for Evidence-Based Policymaking Act 2018* (US).

[5] s. 312 (b).

[6] s. 313.

[7] 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 December 2019. These responsibilities include (but are not limited to):

- serving as an agency champion for — and educator of — agency staff and leaders about evaluation

- 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, coordinating and increasing technical expertise available for evaluation and improving the quality of evaluations and knowledge of evaluation methodology and standards

- 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

- the development of new or improved processes to integrate evaluation findings into agency decision-making and other functions

- management of agency-wide evaluation standards and requirements to ensure the scientific integrity of the agency's evaluation activities

- the use and dissemination of evaluation results throughout the agency and to the public, as appropriate.

Source: Executive Office of the President (2019b s. 290.4).

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 their 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 2020a, p. 68).

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 2020a, p. 69). 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 (2020a, p. 71).

The Magenta Book, published by the **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 (HM Treasury 2018), which assists agencies to appraise policy options (with a focus on

cost-benefit analysis), and the Aqua Book (HM Treasury 2015), which provides guidance on producing (and commissioning) high quality analysis to inform policymaking.

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. The Book 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, utilise an appropriate evaluation approach and methods, and be well executed. 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 Book outlines a 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 2020a, p. 87). 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 focussed

on evaluation planning, evaluation conduct and strategic communication of evaluation results (Government of Canada 2020b).

- In **Slovakia**, a list of recommended courses for evaluators is maintained by the Slovak Value for Money Institute 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 2020a, p. 90).

Some countries also build or promote evaluation competencies into the functional or professional streams of their public service. For example, in the **United Kingdom**, competencies in evaluation and research methods are required for staff of particular levels who are part of the Government Social Research professional stream (Government Social Research (UK) 2019).[8] In addition, an evaluation capability framework (and a self-assessment tool) has been developed for Government analysts (HM Treasury nd, nd). 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.[9]

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 societies (which exist in all OECD member countries) (OECD 2020a, p. 96).

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) nd). 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).

---

[8] The Government Social Research profession is one of 28 recognised professions in the UK Civil Service, with one of the purposes of the profession to 'support the development, implementation, review and evaluation of policy and delivery.' Members of the profession come from a range of backgrounds, including psychology, social policy, geography, and social statistics (GSR nd).

[9] 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) (2019a, p. 10); OECD (2020a, p. 92).

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.

While there are examples of peer review processes being used to promote evaluation quality (including for particular evaluations in **Portugal** and **Germany**) (OECD 2020a, p. 84), 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** Evaluation Standards 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 the past in **Canada,** the Treasury Board Secretariat annually and publicly reported on the Health of Evaluation Function, which provided a government-wide snapshot of evaluation in departments and agencies. These reports included an assessment of the quality of completed evaluations. 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 2020a, p. 87).

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 2020a, p. 106). 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 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 2020a, p. 107).

Approaches for translating evaluation knowledge have also been established in some countries. 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* of 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 area[10] and operates in their own way, but all centres share a common remit that includes knowledge translation (box D.5).

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

---

[10] 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.

(which includes retaining editorial control over research and products), methodological rigour, capacity building and transparency (Cabinet Office (UK) 2018, p. 3).

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).

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 'rate' 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 housed within 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*, who 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.

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 2020a, p. 116).

The OECD found that the use of formal management responses and follow-up procedures for evaluation is relatively uncommon at a departmental level (OECD 2020a, p. 116). One example where such a process does exist is **Canada**, where evaluation reports must be accompanied by a management response and action plan.[11] 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 reports even though there is no formal requirement to do so (OECD 2020a, p. 118).

---

[11] 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).

Requirements for 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 2020a, p. 131).

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 that are targeted specifically at indigenous peoples?

- How do countries approach evaluating impacts and outcomes for indigenous peoples when evaluating mainstream policies 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.[12]

Evaluation of policies and programs affecting indigenous peoples in Canada

The indigenous peoples of Canada comprise of the First Nations, Inuit and Mètis. The relationship between First Nations, Inuit and Mètis 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 2020b). 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

---

[12] These participants included: the Lowitja Institute (sub. 50); the National Aboriginal Community Controlled Health Organisation (NACCHO, sub. 95); and the Queensland Indigenous Family Violence Legal Service (QIFVLS, sub. 25).

Canada was the federal department primarily responsible for meeting Canada's commitments and obligations to indigenous peoples. However, in 2017, it was announced that this department would be 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 Mètis

- *Crown-Indigenous Relations and Northern Affairs Canada*, which seeks to renew the relationship between Canada and First Nations, Inuit and Mètis 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 have only recently been established). At the time of writing this appendix, Indigenous Services Canada had published four evaluations (two of which are discussed in box D.6) (Indigenous Services Canada 2019), 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).

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 constraints in data collected on demand for services meant that the extent to which the program met demand could not be measured.

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, 2020).

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 Aboriginal[13] 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 and Desautels 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'(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 (Jacob and Desautels 2014, p. 62).

When 'mainstream' programs — that is, programs that provide services to all Canadians and not necessarily targeted at First Nations, Inuit or Mètis 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 on 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).

---

[13] 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'.

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)

The nature of relationships between Native American Tribes and the Government of the United States is underpinned by numerous instruments. 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, and meets with about one-third of its grantees each year for the purposes of evaluating the impacts of their grants (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 whose function includes 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 Indian Health Service also maintains a planning, evaluation and research branch. It has developed an evaluation policy that outlines some guidance as to 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 Indian Health Service 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 a number of functions — some of which include: 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 great deal 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 must be applied by New Zealand government agencies when evaluating programs or policies 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 (box D.7) and the extent that they are still drawn upon is not clear.

More recently, Te Arawhiti (the Office for Māori Crown Relations) — established in 2018 — has 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 how to develop an engagement strategy.

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 of 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).

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 capability component and an organisational capability component — that set out different levels of capability maturity and provides 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 functions of Te Puni Kōkiri. Te Puni Kōkiri (or the Ministry of Māori Development) is a government agency with functions that include: leading work towards policy and legislative changes that will improve outcomes for Māori; influencing the work of others by working in partnership and bringing Māori voices to decision makers; and investing with Māori to pursue opportunities (TPK 2020). It also has a legislated monitoring function to assess the 'adequacy' of services delivered to or for Māori (Minister for Māori Development 2019, p. 1).

This monitoring function has used different approaches over time 'to accord with government priorities and the prevailing public management environment' but this function has recently been 'refreshed' (Minister for Māori Development 2019, p. 1). As part of this refresh, Te Puni Kōkiri will monitor Māori wellbeing outcomes and progress across government priorities. Further, Te Puni Kōkiri will also undertake effectiveness reviews. The intent of these reviews is to understand the effectiveness of priority policies, programs and services for Māori and their contribution towards Māori wellbeing (p. 2). The Minister for Māori Development (2019) described effectiveness reviews as:

> … a necessary complement to the previous two levels of monitoring because statistical monitoring alone will not provide an explanation as to *why* specific outcomes have or have not been achieved or *how* effectiveness for Māori could be improved. (pp. 6–7)

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 on 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).

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 of 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).

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).

> ### 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).

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 has been 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' (EvalPartners 2020).