



OECD Recommendation on

# **Public Policy Evaluation**

---

# **Implementation Toolkit**

This paper was approved and declassified by the Public Governance Committee and the Committee of Senior Budget Officials on 25 January 2025.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Revised version April 2025: Two paragraphs have been added to pages 43-44, Annex B

© Layout by Carlos Hinojosa and Clara Aletti

© OECD 2025



Attribution 4.0 International (CC BY 4.0)

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (<https://creativecommons.org/licenses/by/4.0/>).

**Attribution** – you must cite the work.

**Translations** – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of original work should be considered valid.*

**Adaptations** – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

**Third-party material** – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

---

# Abstract

This OECD toolkit on public policy evaluation offers practical guidance for government officials and evaluators seeking to improve their evaluation capacities and systems, by enabling a deeper understanding of their strengths and weaknesses and learning from OECD member country experiences and trends. The toolkit supports the practical implementation of the principles contained in the 2022 OECD Recommendation on Public Policy Evaluation, which is the first international standard aimed at supporting the establishment of robust institutions and practices that promote the use of public policy evaluations. Together, the Recommendation and this accompanying toolkit seek to help governments build a culture of continuous learning and evidence-informed policymaking, potentially leading to more impactful policies and greater trust in government action.

# Table of contents

Executive Summary	5
The purpose of the Toolkit and the Recommendation	6
Pillar 1: Institutionalisation	9
Pillar 2: Quality	16
Pillar 3: Impact	28
Conclusion & additional support	35
References	37
Annex A: Main text of the Recommendation on Public Policy Evaluation	40
Annex B: Self-assessment tool & rubrics	43
Annex C: Evaluation guidelines repertory	58
Annex D: National Evaluation Champions in OECD member countries	64

# Executive Summary

Enhanced public policy evaluation practices have the potential to support evidence-based decision making, accountability, and transparency in governance. Public policy evaluation provides crucial evidence to help policymakers understand what works, for whom, and under what circumstances. When combined with other performance and budget management solutions (e.g. monitoring, performance budgeting and spending reviews), evaluation can contribute to the development of a virtuous cycle of evidence-based policymaking, informed budgetary decision making and fiscally responsible governance. Yet, despite these potential benefits, many governments face challenges in systematically integrating evaluation into their decision-making processes. These difficulties stem partly from limited evaluation cultures and capacities and a lack of evaluation requirements.

The 2022 OECD Recommendation on Public Policy Evaluation offers guidance to countries seeking to effectively build policy evaluation into public governance, organised around three pillars:

- **Institutionalisation of public policy evaluation:** Institutionalising evaluation from a whole-of-government perspective fosters a more formal, co-ordinated, and systematic approach across institutions and levels of government. It creates a robust framework of incentives, responsibilities and accountability of different government, encouraging regular and consistent use of evaluations. By embedding evaluation practices into governance structure, governments can ensure resource efficient and reliable processes.
- **Quality of public policy evaluation:** High quality evaluations are essential to ensuring their relevance for decision making, learning and accountability. Conversely, poor quality evaluations carry the risk of providing unfit evidence, or evidence that is subject to bias and undue influence.
- **Impact of public policy evaluation:** The use of evaluations can help identify which policies work, (and which do not) and why, and thus contribute to practical decision making. Without impactful evaluations and other sources of policy evidence, there is an increased risk of decision making based on instinct rather than facts, potentially leading to ineffective policies. Establishing mechanisms to embed evaluation results in decision-making processes, packaging and presenting evidence from evaluations in a user-friendly way and fostering demand for evaluations can help overcome this challenge.

This toolkit is intended to support countries in their pursuit of the goals set forth in the Recommendation, by providing clear explanations of the Recommendation's concepts and principles, and showcasing relevant initiatives undertaken by OECD Member countries, as well as the trends identified by the 2023 OECD survey on public policy evaluation. At the same time, the toolkit includes a self-assessment tool to facilitate the assessment of governments' existing evaluation capacities, along with their strengths and potential areas for improvement. Such an assessment can serve as a starting point for the development of an evaluation capacity improvement roadmap.

# The purpose of the Toolkit & the Recommendation

## Why is it important to enhance public policy evaluation practices and institutions?

Public policy evaluation brings an objective understanding of what works, why, for whom, and under what circumstances. It provides crucial evidence to support decision makers, enhancing the transparency and accountability of government actions (OECD, 2020[1]). Indeed, evaluation represents an important building block for the development of evidence-informed policymaking and sound public spending, and fostering a culture of learning and continuous improvement. By developing this culture, governments can also increase the levels of trust and understanding by citizens in the actions they are undertaking.

Yet, ensuring a widespread use of evaluations as a basis for effective and informed policy and decision making remains a challenge for many governments. Buy-in and political support for evaluation practices often requires an important paradigm shift throughout government and civil servants which does not always take place organically. As such, there is a need for guidance and spaces for government practitioners to better understand the strengths and weaknesses of their own systems, as well as the existing knowledge and experiences available to enhance them.

**‘Public policy evaluation’ refers to the structured and evidence-based assessment of the design, implementation or results of a planned, ongoing or completed public intervention. It assesses the relevance, coherence, efficiency, effectiveness, impact and/or sustainability of a policy based on its objectives (OECD, 2022[2]).** This broad definition is purposely meant to cover a broad range of evaluation approaches from a methodological (e.g. theory-based, experimental), policy cycle (e.g. *ex ante*, *ex post*), and instrument perspective (e.g. policies, programmes, legislation). It also acknowledges the diversity of definitions and approaches to evaluation used in different country contexts. Toolkit users should reflect on the types of evaluations they wish to apply the toolkit to and adapt its use accordingly.

## What is the OECD Recommendation on Public Policy Evaluation?

The OECD and its member countries developed a systemic approach that examines individual country evaluative practices and how these can come together to ensure that evaluation becomes fully integrated into the policy cycle. This approach is embodied in the 2022 OECD Recommendation on Public Policy Evaluation. The Recommendation represents the first international standard aimed at driving the establishment of robust institutions and practices that promote the use of public policy evaluations. It is structured around three mutually reinforcing pillars and seven dimensions (see page 7 & Annex A).

The Recommendation complements other OECD legal instruments that directly or indirectly address policy evaluation such as the Recommendation on Improving the Quality of Government Regulation (OECD/LEGAL/0278) or the Recommendation on Regulatory Policy and Governance (OECD/LEGAL/0390). It also complements the OECD’s longstanding work on evaluation of development co-operation including the OECD Glossary of Key Terms in Evaluation and Results-Based Management for Sustainable Development.

## What is the purpose of this Toolkit and who should use it?

The Toolkit aims to explain and operationalise the goals, principles and concepts outlined in the Recommendation and offer practical implementation advice and examples. It is designed for a range of users, including government officials in centres of government (e.g., Ministries of Finance, Executive offices) responsible for their country's evaluation policy, evaluation officers in line ministries, legislative bodies or national agencies, evaluation officers in regional or local governments, regulators, or anyone wishing to enhance the evaluation capacities of their organisation. The Toolkit can be used to frame an overarching reflection on existing capacities and potential areas for improvement at the national level, within specific branches of government or line ministries, or at subnational level. It may also be used to explore and assess specific aspects underpinning evaluation capacity, such as institutionalisation, quality, or impact of evaluations.

## How can this Toolkit be used?

There is no one-size-fits all approach to using this Toolkit and its use can be adapted to the needs and interests of each user. Users may wish to learn about one or more specific pillars or dimensions of the Recommendation, identify good practices implemented by other countries as well as OECD-wide trends (**inspire**), understand the concrete measures which can contribute to achieving the goals of the Recommendation (**act**), access existing learning material (**learn**), or alternatively conduct a self-assessment to learn more about their existing capacities and identify opportunities for improvement (**assess**). For users interested in the self-assessment component, Annex B contains a dedicated and detailed self- assessment tool.

The toolkit is structured following the three pillars of the Recommendation.

## The 4 intended uses of the Toolkit



Take stock of the current situation with regard to the principles and standards promoted by the Recommendation and identify opportunities for improvement.



Identify trends and good practices from other OECD Members, which may act as a reference for the development of local solutions.



Identify potential practical solutions and tools to action the goals and principles of the Recommendation.



Access and identify existing resources developed by other countries and organisations allowing to learn more about specific issues and potential methods / solutions.

# The **three pillars & seven dimensions** of the OECD Recommendation on Public Policy Evaluation

## **Pillar 1: Institutionalisation of Public Policy Evaluation from a whole of government perspective**



Dimension 1:  
Execution & Agency

Dimension 2:  
Learning & Accountability  
Culture

## **Pillar 2: Quality of Public Policy Evaluation**



Dimension 3:  
Design, Planning  
& Management

Dimension 4:  
Quality & Standards

Dimension 5:  
Institutional Capacities  
& Skills

## **Pillar 3: Impact of Public Policy Evaluation**



Dimension 6:  
Use in decision-making

Dimension 7:  
Accessibility & Dissemination  
of Findings & Results

# Pillar 1: Institutionalisation of Public Policy Evaluation from a whole of government perspective

## Why does institutionalisation matter?

Institutionalising policy evaluation within different institutions and levels of government can contribute to integrating isolated and unplanned evaluation efforts into more formal and systematic approaches. Further, institutionalisation can create the adequate framework of incentives, responsibilities and accountability of different government agents to carry out and make use of evaluations more regularly. In return, there may be higher frequency and consistency of evaluations, and thus comparability and reliability of results across time and institutions.

Yet, institutionalising evaluation throughout government requires a significant paradigm shift which may prove challenging. In addition, there is no one-size-fits-all solution to establishing government-wide institutional frameworks for policy evaluation. While some countries have adopted highly formalistic approaches as reflected in the adoption of high-level legislation on the issue, others rely on more informal yet-generalised practices throughout government and civil servants. OECD member countries are at different stages of the “institutionalisation spectrum”, with some yet to overcome key barriers to fully achieving this goal. While OECD Members have increasingly adopted “evaluation champions” or central evaluation units responsible for coordinating policy evaluation across government; the roles, responsibilities and resources allocated to other stakeholders inside and outside of government vary considerably.



Inspire

## What the OECD survey on Public Policy Evaluation says about institutionalisation in member countries

84% put in place **at least one evaluation champion**

61% have adopted one or several **formal definitions of evaluation**

48% have **knowledge brokers** which are involved in policy evaluation

45% have **Supreme Audit institutions with a mandate** to conduct policy evaluations

35% of **legislative bodies** have an internal body responsible for conducting policy evaluations

# Dimension 1: Execution and Agency

## What does the Recommendation say?

Countries should conduct and use evaluations across government ensuring that they are carried-out in a systematic manner and that their results are used in policy and budgetary decision making.

## What is the aim?

**Ensuring institutions have a clear and shared understanding of their roles and responsibilities in planning, conducting and using evaluations.**

Defining and assigning institutional responsibilities for conducting and using policy evaluations is key to ensuring evaluation is cemented into the work and missions of different government institutions, and that adequate time and resources are dedicated to it. This also minimises the risk of overlap and contributes to institutions knowing when and how to conduct evaluations so that their results can feed into decision making.

**Transitioning from a fragmented approach of evaluation to a more formal and systematic one, increasing the consistency and reliability of results.**

Without formal guidance or clear guidelines on who does what and when, evaluations and other related exercises (e.g. monitoring, spending reviews), often take place in isolation and end up being far removed from policy decisions. Strengthening the regularity and co-ordination of evaluation exercises can contribute to improving the reliability of results, and to the achievement of economies of scale in the delivery of evaluation functions.

**Providing leadership and guidance in support of the delivery of evaluation responsibilities and the use of evaluations across government and within individual organisations.**

The creation of an “evaluation champion” or networks of evaluation focal points can improve the coordination of policy evaluations across institutions and support the system by promoting the quality and use of evaluations. The allocation of such a central role to a new or existing agency, as well as the type of role and location in government, can send a strong message regarding the importance given to evaluation.

**Evaluation champions** are actors or institutions inside the government that have a cross government mandate related to policy evaluation (See Annex D). Their exact responsibility can vary across countries, but often includes co-ordinating cross-sectoral policy evaluations, and sharing best practices to promote their quality and use (e.g. proposing standards for evaluation).



- Are institutional responsibilities for conducting evaluations clearly defined and assigned throughout government?
- Are evaluations conducted in a co-ordinated manner and according to best practices thanks to the support and advice provided by one or several evaluation champions?
- Do evaluations support and build on other complimentary performance management exercises such as monitoring, spending reviews and internal audits?



## What initiatives have OECD member countries introduced?



### **Legally embedding evaluation and establishing clear evaluation mandates and responsibilities: The Japanese Government Policy Evaluations Act and Spain's 2022 Law on the Institutionalisation of Evaluation**

The Japanese government embedded evaluation across its administration as early as 2001 with the Government Policy Evaluations Act No. 86. The Act provides a framework to assess the impact of proposed legislation, incorporating results into planning, and defines the roles of administrative organs responsible for these assessments. It mandates guidelines on ex ante regulatory impact assessment and mechanisms for reporting results to the legislative body (the Diet) and the public (Government of Japan, 2001[3]). In 2017, the guidelines were further updated to include ex post evaluation.

In Spain, Law 27/2022 institutionalised policy evaluation across government. It defines clear mandates and institutional responsibilities on policy evaluation across government. In addition, it creates the State Agency for the Evaluation of Public Policies to champion, supervise, co-ordinate, and promote evaluations. The law mandates line ministries to create evaluation co-ordination units and assigns the Independent Fiscal Institution to conduct ex post evaluations through spending reviews (Head of State, 2022[4]). Clarifying the role of different actors in a single piece of legislation can considerably help in ensuring that there is a common understanding of the different functions, responsibilities and expectations.



### **National evaluation champions: the Ministry of Finance in the Netherlands and the newly created PlanAPP in Portugal**

In both the Netherlands and Portugal, specialised governmental units act as national “evaluation champions” playing central roles in policy evaluation. In the Netherlands, the Strategic Analysis Unit (SAU) within the Ministry of Finance leads the evaluation efforts. It co-ordinates policy evaluation across public administration, managing the Strategic Evaluation Agenda and providing advisory services, best practices, and central guidelines like the Policy Evaluation Toolbox. The SAU oversees the entire evaluation cycle, from planning to disseminating results, and strengthens monitoring and evaluation (M&E) capacity through training developed by the National Academy of Finance. It also shares evaluation outcomes via a central evaluation platform.

Similarly, in Portugal, PlanAPP, established in 2021, serves as a centre of expertise for public policy evaluation. It promotes best practices and the effective use of evaluation results in public policy development. PlanAPP co-ordinates RePLAN, an inter-ministerial network focused on planning, foresight, and evaluation. This network is in the early stages of establishing a national evaluation agenda and common guidelines to enhance public administration's evaluation efforts. Both the SAU and PlanAPP play pivotal roles in improving policy evaluation processes and outcomes in their respective countries.



- Define and assign institutional evaluation responsibilities through for instance, the development of dedicated legal and / or policy frameworks.
- Adopt a shared and formal definition of evaluation which can be included in guidance on evaluation, legislation, or evaluation policies.
- Create and / or empower evaluation champions at the cross-government level, as well as within individual government institutions.



## Keep in mind

- Provide evaluation champions with the authority and resources to reward, incentivise and influence line ministries and other government stakeholders involved in evaluation, to conduct evaluations.
- Where the evaluation champion sits in government can be as important as having a champion. Champions should be able to access to decision-making processes and work across government to avoid isolation. In OECD member countries, evaluation champions are generally located in centres of government, ministries of finance, or exist as autonomous agencies.
- Equip evaluation champions with skills including political acumen, strong communication and networking and co-ordination skills.

- Evaluation guidelines can support governmental actors define coherent policy evaluation processes and methods. However, avoid overly-detailed and/or prescriptive guidance and duplication, as this can lead to confusion and redundancy (Guidelines are further discussed in Dimension 4: Quality and Standards).



- Lazaro, B. (2015[5]), Comparative study on the institutionalisation of evaluation in Europe and Latin America. This report provides insight on the development of public policy evaluation systems in European and Latin American countries (<https://tinyurl.com/LazaroB>).
- OECD (2020[6]), Improving Governance with Policy Evaluation: Lessons from Country Experiences, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/89b1577d-en>. The report provides an overview of the institutional frameworks that enable policy evaluation, and promote its quality and use.

## Dimension 2: Learning and Accountability Culture

### What does the Recommendation say?

Countries should foster a culture of learning and accountability by promoting demand for, and ownership of, evaluations within and beyond the executive.

### What is the aim?

**Ensuring that the use of and demand for evaluations exists across all branches, levels and functions of government, and endures political changes.**

While the institutionalisation of evaluation may initially take hold at the centre of government, it's important for all government institutions and functions outside of the executive, at all levels, to eventually support and use evaluations. Non-governmental evaluation societies as well as other non-governmental evaluators (e.g. academic researchers) play a key role in developing such an evaluation culture, which is key to ensuring a sustainable and resilient institutionalisation of evaluation.

**Co-ordinating the supply and the demand for evidence throughout government with the support of a community of “knowledge brokers”.**

Institutionalisation of policy evaluation and evidence-informed policymaking requires a specific set of actors known as “knowledge brokers”. Knowledge brokers play an important role in bridging the divide between producers and users of evidence. They carry out a range of functions, including that of ensuring the appropriate evidence makes its way to policymakers in a timely, attractive, relevant, and accessible manner.

**Offering the opportunity for legislative bodies to review and discuss policy evaluations.**

While the focus of institutionalisation efforts is often on the development of evaluation functions, rules and guidelines within executive branches of government, legislative bodies also have a key role to play in commissioning, reviewing and discussing policy evaluations

**Knowledge brokers** are individuals or organisations that strengthen linkages between evidence users (e.g., decision makers, policymakers) and suppliers (e.g. evaluators, researchers). They can take on a variety of forms, ranging from individual professionals (such as Government chief science advisors, or ministerial advisors) to dedicated organisations. In terms of institutions, some are specifically connected to knowledge producers, such as brokering units within academic institutions (OECD, 2020[6]). Knowledge brokers can be located both inside government and at arm's length and in some cases, this function can be undertaken by evaluators or evaluation co-ordinators.



- Are evaluations regularly conducted and reviewed outside of the executive branch of government?
- Do non-governmental institutions and stakeholders play a role in strengthening the country's evaluation culture?
- Do knowledge brokers exist and play a role in strengthening the relationship between evidence from evaluations and its users, including citizens?
- Does the legislative branch use the results of evaluations to help shape new or revised legislation?



## What initiatives have OECD member countries introduced?

### **Policy evaluation outside of the executive: the role of the French National Assembly and Supreme Audit Institution**

The French Supreme Audit Institution (Cour des Comptes) and the National Assembly (Assemblée Nationale) have been invested by the Constitution with important evaluation functions. The Cour des Comptes is mandated to assist the National Assembly in evaluating public policies (Art. 47-2) and it regularly conducts policy evaluations at the request of the Assembly or at its own initiative. In collaboration with the National Assembly, it also organises the national evaluation open days bringing together evaluators across the French public administration and academia. The National Assembly also has the power to conduct and commission its own evaluations (Art 24) (Assemblée Nationale, 2024[7]). To do so it has established the Committee on Public Policy Evaluation and Control chaired by the President of the National Assembly and composed of thirty-six members. In addition, since 2018, the Assembly also conducts a systematic discussion of evaluations defined by an annual evaluation programme. Each year between May and June all ministries present the findings of the defined evaluations and are questioned on the performance of the public policies for which they are responsible (Commission des finances, de l'économie générale, 2023[8]).

### **The Swedish National Audit Office and its role in supporting an evaluation culture in Sweden**

The Swedish National Audit Office is an important institution when it comes to the development of a culture for evidence and evaluation in Sweden. The Audit Office conducts evaluations itself as well as reviews of the country's evaluation system. These evaluations are often shared and discussed with the Parliament. In 2022, the Audit Office published an analysis of the Swedish evaluation system and approach to evidence-informed policymaking which analyses 25 reforms introduced between 2000 and 2017 (Swedish National Audit Office, 2022[9]). More recently the Audit Office reviewed the Swedish development aid monitoring and evaluation system providing recommendations to strengthen evaluation practices in this area (Swedish National Audit Office, 2024[10]).

### **Brokering the supply and demand for knowledge in policymaking: the What Works Network in the United Kingdom**

In the United Kingdom, the What Works Network plays an important support function in making evaluations more impactful and facilitating their use by policy and decision makers. This network, established in 2013, is composed of 12 different independent knowledge centres (more precisely, nine independent centres and three affiliate members) which can support decision makers and practitioners with access to the evidence of "what works" when it comes to policy design and implementation, mainly by producing evidence reviews and syntheses. Particularly on policy evaluation, the organisations of the network provide recommendations to government evaluators and suggestions while assessing and synthesising the most recent evidence in several policy sectors (Evaluation Task Force, 2024[11]). The network is managed by the UK government's Evaluation Task Force – the country's evaluation champion (Evaluation Task Force, 2024[11]) which ensures that government internal evaluators can benefit from the expertise of the different knowledge centres.



- Promote the “knowledge brokerage” function by creating dedicated structures or ensuring existing ones undertake this role.
- Create spaces and mechanisms for legislative bodies to discuss and review evaluation results.
- Promote the development of non-governmental evaluation societies.
- Support the involvement, as relevant, of Supreme Audit Institutions in providing oversight of evaluation practices, and conducting audits which are complementary to evaluations.
- Build synergies with local governments (e.g. exchanging data, co-ordinating evaluations, using common standards) and their evaluation efforts to expand the evidence-base on the effectiveness and impact of regional / local policies



## Keep in mind

- Encourage a learning-oriented approach in evaluation systems instead on focusing exclusively on control or accountability.
- Address the fear of failure and create an environment where unexpected results stemming from evaluations are viewed as opportunities for learning and introduction of meaningful change, rather than failures to be punished. Further, evaluations should pro-actively scope for unexpected results (positive and negative).



- INTOSAI (2019[12]), GUID 9020: Evaluation of Public Policies. These guidelines aim to support Supreme Audit Institutions in the delivery of evaluation activities and encourage the community of auditors to enter the field of evaluation of public policies in order to help them to conduct evaluation properly, scientifically and independently. They are developed by the International Organisation of Supreme Audit Institutions (<https://tinyurl.com/INTOSAIGUID>).
- Speer, S., Pattyn, V. and De Peuter B. (2015[13]), “The Growing Role of Evaluation in Parliaments: Holding Governments Accountable”: The paper provides an analysis of written questions and interrogations in Germany and Flanders’ Parliaments, providing an insight into how parliamentarians view and use evaluations (<https://tinyurl.com/SpeerPattynDePeuter>)

# Pillar 2: Quality of Public Policy Evaluation

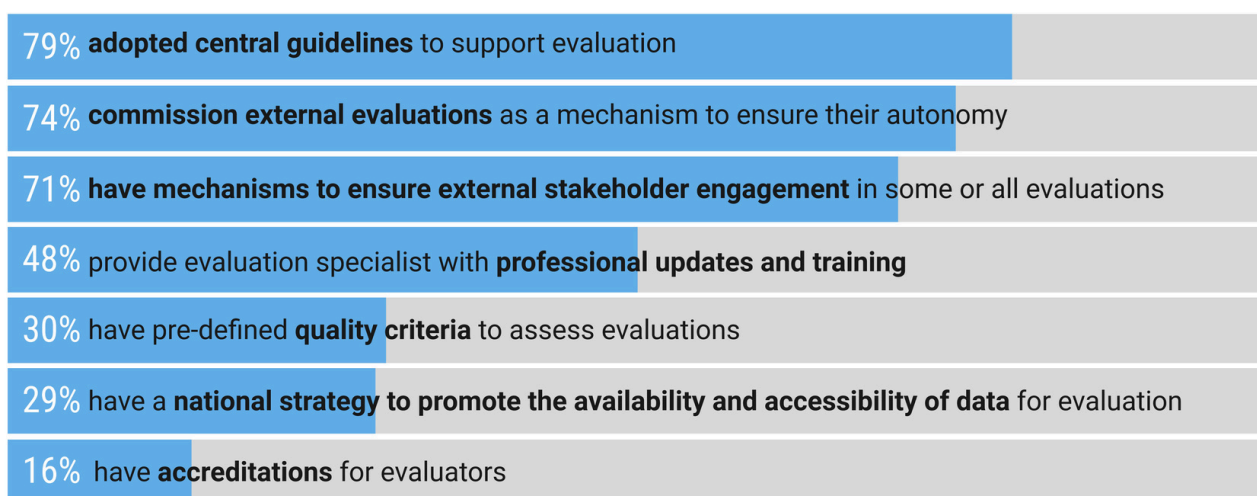
## Why does the quality of public policy evaluation matter?

Ensuring evaluations meet high quality standards is essential to increase their use for decision-making, learning and accountability purposes. Conversely, poor quality evaluations carry the risk of providing unfit evidence, or evidence that is subject to bias and undue influence (OECD, 2020b). Quality refers to the technical rigour with which evaluations are conducted to avoid bias and ensure reliability, as well as to the governance processes used to conduct them and make use of their results in policymaking. Governments can use a range of solutions to ensure they produce quality evaluations and reliable evidence, including the adoption of quality standards and principles of good evaluation governance. Further, well-developed skills are essential to conducting high quality evaluations, placing evaluators and their skills at the heart of successful evaluation systems.

OECD member countries have increased efforts to perform high-quality evaluations, as reflected by the fact that 79% have adopted central evaluation guidelines. However, evaluation quality remains one of the most significant challenges to the development of strong evaluation cultures and the use of evaluation evidence in OECD member countries. This is illustrated by the limited human resources and poor quality of evidence, limiting the ability of promoting evaluations across government in OECD countries. In addition, only a small fraction of OECD member countries have adopted criteria to assess the quality of the evaluation across government.



### What the OECD survey on Public Policy Evaluation says about quality in member countries



## Dimension 3: Design, Planning and Management

### What does the Recommendation say?

Countries should actively plan, design, and manage evaluations so that they are timely and proportionate to the intended objectives, taking into account the needs of the primary users and the types of intended uses, and ensuring that results can be trusted by stakeholders

### What is the aim?

#### Ensuring that evaluations are use and user-focused, timely, and fit for purpose

Evaluations must be planned and carried out in anticipation of the processes and decisions they are meant to support. Further they should be designed while keeping in mind the intended uses and users of the evidence they generate, as well as to guarantee the production of relevant evidence for potential future policy decisions.

#### Selecting the adequate and relevant methods to carry out the evaluation, based on the objectives being pursued and the type of intervention being assessed.

Evaluations can be designed in multiple ways, to explore different evaluation questions and criteria, and potentially generate evidence through a range of different data collection and analysis methods. Adequately designing evaluations and selecting the relevant methods to conduct them is key to ensuring robust and reliable results.

#### Promoting the involvement of the relevant stakeholders in key steps of the evaluation, reducing potential bias and increasing the likelihood of impact upon completion.

Stakeholder involvement in evaluations is necessary to ensure the production of quality results, as well as to increase the likelihood for buy-in and support for the recommendations presented. However, involving stakeholders can be time and resource consuming, and may also increase the complexity of the evaluation process. Strong evaluation systems thus encourage and facilitate stakeholder involvement in evaluations, without limiting their timely delivery and execution..

**Evaluation clauses** refer to a clause expressed in a legal act requiring the evaluation of said act. They are important to ensure that the act is subject to some sort of review to ensure they remain effective and appropriate over time (OECD, 2021[14]).



- Are evaluations planned well in advance and delivered in a timely fashion, and are provisions for evaluations (e.g. evaluation clauses) built into public interventions from the start?
- Are relevant stakeholders engaged and mobilised throughout the evaluation process to create ownership for change and trust in evaluation results?
- Are there procedures in place to select the most relevant evaluation methods based on the type of evaluation being pursued and their intended use?



## What initiatives have OECD member countries introduced?



### Embedding ex post evaluation in the policy cycle: the European Commission's use of evaluation clauses and "evaluate first" principle

The European Commission employs several strategies to embed ex post evaluation in the policy cycle in support of evidence-based policymaking. One key approach is the incorporation of evaluation clauses in legislative proposals, which explicitly outline the timing, accountability, content, and format of future evaluations. This proactive measure ensures that evaluation is an integral part of the policy cycle from the outset. The Commission has introduced the 'evaluate first' principle, which requires that evaluations be conducted before new initiatives are proposed. By prioritising evaluation, the Commission can assess the cost-effectiveness, coherence, relevance and EU added value of existing policies and identify areas for improvement, ultimately informing the development of new initiatives. By evaluating existing policies before proposing new ones, the Commission can identify what works and what doesn't, informing policy decisions and adjustments.

### Strategically planning evaluations in advance: Italy's three-year evaluation plans

Italy introduced a new government-wide framework for policy evaluation, integrating the practice into the national budget cycle and enhancing ministries' capacity to plan and conduct evaluations. Starting in 2024, each ministry must develop an evaluation plan outlining policies to be evaluated, evaluation questions, timing, and costs (Ministry of Economic and Finance, 2023[15]). Coordinated by the Ministry of Finance, the system allows autonomy in designing evaluations but requires sharing results and methods. The Ministry's Department of General Accounting, through a Scientific Committee, provides support, guidelines, and technical advice on methods, timelines, policy selection, and stakeholder engagement.



### Ensuring stakeholder engagement in evaluations: Costa Rica's stakeholder engagement guidelines

The Costa Rican government has adopted specific mechanisms to ensure stakeholder engagement in evaluation processes. In this context, the Ministry of Planning and Economic Policy (MIDEPLAN) developed guidelines on stakeholder engagement and participatory evaluations in support of the evaluation work carried out by line ministries (Ministry of National Planning and Economic Policy, 2019[16]). The guidelines themselves were developed through a participatory approach, through the collection of input from stakeholders in the country's six planning regions. The guidelines are structured around three main topics including an introduction to the key concepts relating to participatory evaluation, a presentation of suggested practices to enhance the participatory nature of evaluations, and practical tips for the development of the evaluation process.



- Implement evaluation planning mechanisms and tools such as evaluation plans. Evaluation agendas and plans are valuable tools to anticipate and think early about evaluation needs, and prioritise and conduct evaluations in the areas where they are most needed.
- Include evaluation incentives or obligations in policy initiatives, such as evaluation clauses
- Involve evaluation users in the design phase, and design evaluations based on future intended uses.
- Adopt participatory approaches to evaluation



## Keep in mind

- Evaluation plans and timetables should be aligned with key decision-making processes (e.g. the budget process, government programmes) to increase the likelihood of impact (Evaluation impact is further discussed Dimension 6: Use in decision making). However, evaluation plans should include a level of flexibility to evaluate time sensitive interventions (e.g. COVID responses).
- Engage stakeholders clearly and purposefully in relevant stages of evaluations through clear mechanisms and well-defined purposes.
- Balance stakeholder engagement with the time and resources available to avoid overburdening the evaluation process.



- Preskill, H. and N. Jones (2009[17]). A Practical Guide for Engaging Stakeholders in Developing Evaluation Questions: This guide was designed by the Robert Wood Johnson Foundation to help evaluators in engaging with stakeholders. In particular, it provides a five-step process to design evaluation questions with stakeholders. (<https://tinyurl.com/PreskillJonesRWJF>)
- OECD (2020[18]), Reviewing the Stock of Regulation, OECD Best Practice Principles for Regulatory Policy: This paper identifies good practice principles drawn from OECD member countries to strengthen the review of existing regulation. The paper offers several examples illustrating the implementation of these principles, covering the governance, the methodology, and the importance of public consultation for the review of regulation. (<https://doi.org/10.1787/1a8f33bc-en>)
- OECD (2021[19]), Applying Evaluation Criteria Thoughtfully: This guidance aims to help evaluators and others to better understand the evaluation criteria of relevance, coherence, effectiveness, efficiency, impact, and sustainability, and improve their use to determine the merit or worth of various interventions. (<https://doi.org/10.1787/543e84ed-en>)
- Patton, M. Q. (2002[20]), Utilization-focused evaluation checklist: The checklists can be used throughout the evaluation process. They distil and clarify relevant elements to keep in mind when planning and managing evaluations. They can be particularly useful during the initial stages to ensure a good overview of the different actions to undertake. The checklists combine practitioner experience, theory, principles, and research. (<https://tinyurl.com/PattonMQ>)

## Dimension 4: Quality and Standards

### What does the Recommendation say?

Countries should establish quality standards and mechanisms for evaluations to generate robust and credible evaluation results that can be trusted and used with confidence.

### What is the aim?

#### **Ensuring the quality and reliability of evaluation results by using state of the art, robust and relevant methods, and conducting quality assurance processes**

By issuing centralised guidelines for evaluation, governments can support evaluators select methods and approaches which are fit for purpose, and ensure the use of relevant, cost-efficient, and cutting-edge approaches. Central guidelines on quality standards and methods of assessment can also mitigate the risk of delivering poor quality evaluations.

#### **Guaranteeing the independence and objectivity of evaluation exercises and results and mitigate any potential risk of bias**

Evaluations may generate different risks of bias based on the proximity which may exist between stakeholders responsible for designing and delivering policy initiatives, and those responsible for evaluating them. While certain methodological solutions may reduce this risk (e.g., triangulation of evidence, and participatory methods), adequate evaluation and evidence governance mechanisms are also necessary.

#### **Safeguarding the dignity, safety and privacy of stakeholders involved in the delivery of evaluations**

Conducting evaluations may raise ethical considerations, particularly as evaluators interact with, gain access to and draw information from policy stakeholders (e.g., personal data from beneficiaries of interventions). Measures should be implemented to ensure that evaluations are carried out according to the highest ethical standards, relevant regulations (e.g. Human Research Ethics Regulations), or standards (e.g. United Nations Declaration on the Rights of Indigenous People), and that evaluators are aware of and able to manage trade-offs between quality, reliability and ethics.

**Evaluation guidelines** refer to a document providing directions on how, why and when to conduct policy evaluation, for instance in terms of what methods to use, which stakeholders to consult, how to manage the process or how to use evaluation results. Annex C of the toolkit includes a list of national and European evaluation guidelines.



- Are methodological best practices being used when designing evaluations, conducting data collection and delivering analyses?
- Are integrity, independence and cultural sensitivity being ensured in evaluations through the use of professional and ethical standards for evaluators?
- Are internal and external evaluations being conducted with sufficient levels of autonomy?
- Are evaluations assessed systematically against pre-defined quality criteria to ensure the validity and robustness of their results?



## What initiatives have OECD member countries introduced?



### **Providing comprehensive guidelines for policy evaluation: the Australian Evaluation Toolkit**

The Australian Government's updated 2023 Commonwealth Evaluation Policy and Toolkit provides access to simple and practical guidance on how to plan and conduct evaluations. The toolkit covers all the steps of an evaluation starting from what it is and why it is done, to how its results should be communicated (Australian Centre for Evaluation, 2023[21]). It also includes templates, tools and additional resources showcasing best evaluation practices. The toolkit was intentionally designed to be concise and suitable for non-technical audiences. Further, it's publicly accessible on-line with the aim of improving navigability and facilitating regular updates with additional resources and supporting material.



### **Defining quality standards and practices for evaluation: the United States' Evidence Act**

The 2018 Foundations for Evidence-Based Policymaking Act (the Evidence Act) has provided significant thrust to policy evaluations in the United States, by recognising programme evaluation as an essential component of federal evidence building (Congress, 2018[22]). Alongside this effort to further systematise programme evaluation across the government, in 2020 the Office of Management and Budget developed high-level standards to be followed by federal agencies when conducting policy evaluations. The standards are structured around five core areas: relevance and utility, rigor, independence and objectivity, transparency and ethics; and they are accompanied by ten leading evaluation practices to support them (Office of Management and Budget, 2020[23]).



### **Ensuring evaluators' adherence to ethical standards: Czechia's evaluator's code of ethics**

The Czech Evaluation Society developed the Evaluator's Code of Ethics in 2011. The code was adopted with the aim of improving quality and reputation evaluation practices at the national context, as well as to enhance their professionalisation and standardisation. The code was developed through extensive consultations with evaluators and evaluation stakeholders nationally. It covers 21 principles which are in turn organised according to three dimensions: expertise, integrity, and responsibility. The Code is currently used by the government in the context of outsourced evaluations of programmes funded by the European Union's Structural Funds.



## What initiatives have OECD member countries introduced?

### **Assessing the quality of evaluations: Portugal's Evaluation Report Quality**

### **Assessment Grid and the European Commission's Regulatory Scrutiny Board**

The Agency for Development and Cohesion in Portugal (AD&C) has developed several tools to support the development of quality evaluations and their assessment. Among these tools is the Quality Assessment Grid aimed at ensuring the quality of evaluations of programmes funded by the European Union's Structural Funds in Portugal. The tool is particularly relevant given that most of these evaluations are outsourced to external evaluators. The grid contains a standardised assessment framework applicable to different phases of evaluation (inception, interim, and final reports), ensuring consistency in the use of quality criteria while allowing for a certain level of flexibility to assess evaluations using different methodologies. The assessment grid primarily focuses on the methodological adequacy of the evaluation, the rigour in the application of methods and adherence to standards, and the transparency and rigor in the analysis.

In the context of European policymaking, the European Commission's Regulatory Scrutiny Board is an independent body providing central quality control and support for Commission impact assessments and evaluations at early stages of the legislative process. The Board reviews and issues opinions and recommendations on all the Commission's draft impact assessments, fitness checks, and on major evaluations of existing legislation. The Board also provides cross-cutting advice on better regulation policy to the Commission's Secretariat-General (European Commission, n.d.[24]).



- Adopt evaluation, quality and evidence standards through for instance, evaluation guidelines, or legal / policy frameworks. Standards may include technical standards relating to the methodological soundness of evaluations, and good governance standards relating to ensuring the independence, impartiality and transparency of the evaluation process.
- Implement evaluation quality assessment and control procedures, such as peer reviews or metaevaluations.
- Develop evaluation guidelines such as evaluation whitebooks.
- Implement professional and ethical standards for evaluators inside and outside of government.
- Conduct audits of evaluation systems, through for instance, Supreme Audit Institutions.



## Keep in mind

- Ensure evaluation guidelines are updated to reflect innovative and state-of-the-art evaluation approaches.
- Over-reliance on commissioning external evaluations to guarantee their independence can erode internal capacity to conduct evaluations. Balance external expertise with efforts to build and maintain internal evaluation skills and resources.
- Independence should be sought not only in the conduct of evaluations, but also in their reporting and use. The use of conflict-of-interest forms for internal and external evaluators can help safeguard the integrity and prevent bias of external evaluation processes.
- Procurement processes to outsource evaluations might favour lower-quality offers due to cost considerations. Strive to balance cost with quality to ensure that evaluations meet the required standards and provide valuable insights.



- Australasian Evaluation Society, Ethical Guidelines: Guidelines for ethical behaviour and decision making in evaluation intended to foster continuing improvement in the theory, practice and use of evaluation by stimulating awareness and discussion of ethical issues. The Guidelines refer to the ethical principles to be observed at three of the main stages of evaluation: commissioning and preparing, conducting, and reporting. (<https://tinyurl.com/AESethics>).
- Nutley, S. Powell, A. and Davies H. (2013[25]), What counts as good evidence? Alliance for Useful Evidence: The paper reviews the ways of assessing and labelling evidence from research. It notably discusses the risks of using a hierarchy of evidence and explores alternative approaches which may be better suited for a variety of evidence-informed decision-making processes. (<https://tinyurl.com/NutleyPowellDavies>)
- OECD (2010[26]), Quality Standards for Development Evaluation: With a view to improving the quality of development evaluation processes and products, and facilitating collaboration, this reference guide lays out standards for each phase of a typical evaluation process: from defining purpose, to planning, designing, implementing, reporting, and learning from and using evaluation results. (<https://doi.org/10.1787/9789264083905-en>)
- Revillard, A (2023[27]), Policy Evaluation: Methods and Approaches: A compendium presenting 24 qualitative, quantitative or mixed methods or approaches in evaluation. The methods are presented in a didactic and illustrated manner, based on a common series of questions that facilitate their comparison. (<https://tinyurl.com/Revillard>)
- Varazzani, C. et al. (2023[28]), Seven routes to experimentation in policymaking: A guide to applied behavioural science methods: This working paper proposes guidelines and a visual roadmap, including five key questions to guide the selection of appropriate methods, and seven different experimental or non-experimental methods to choose from. The aim is to provide a simple and practical framework for policymakers to choose the most fit-for-purpose method for their policy context. (<https://tinyurl.com/Varazzani>)

## Dimension 5: Institutional Capacities and Skills

### What does the Recommendation say?

Countries should develop institutional skills and capacities to conduct, commission and use evaluations effectively and in a credible manner.

### What is the aim?

**Ensuring the availability of the necessary skills and competencies to manage and carry out evaluations and responsibilities across government, generating sufficient demand and supply of quality evaluations and evidence.**

Individuals who possess the right skills and competencies are more likely to produce or commission high quality user-focused evaluations, and increase the supply and demand for evaluations. As such, there is a need for governments to attract the necessary talent to conduct evaluations, as well to as upskill and reskill existing staff at all levels to become more versed in the design, implementation and use of evaluation methods.

**Generating and ensuring access to the necessary data to conduct robust evaluations, while respecting key privacy and confidentiality principles.**

The availability of high-quality data in a timely, accessible, disaggregated, inter-operable and re-usable format is crucial to performing quality and reliable evaluations. Governments must ensure that the necessary data is available for evaluation purposes – through for instance the development of strong monitoring systems -, while managing and mitigating the potential risks associated with the publication of large volumes of sensitive data.

**Securing and providing the necessary human and financial resources to carry out evaluation functions and produce quality evaluations.**

Commissioning, conducting and using evaluation requires resources, which is why governments must adopt specific measures such as earmarking of funds or specific budget lines to finance the delivery of evaluation functions. This must be done for central evaluation units as well as at the level of individual ministries or agencies.

**Capacities** are an organisation's resources, both in terms of people and their skills and capacities, and of budget, dedicated to the activity of evaluations.



- Are employee skills for policy evaluation being acquired, retained and developed at the necessary pace and levels?
- Is the availability and accessibility of data for research and evaluation adequately regulated, promoted and recognised?
- Do evaluation units and teams have access to the necessary human and financial resources to carry out all evaluation functions?



## What initiatives have OECD member countries introduced?

### **Strengthening the ability to evaluate effectively: the innovative learning and development model of the United Kingdom's Evaluation Task Force**

The United Kingdom's Evaluation Task Force (ETF) ran a five-day evaluation academy in March 2023 to train and upskill 40 future evaluation trainers from 14 departments of the UK government in evaluation methods and other soft skills (e.g., evaluation advocacy and communicating evidence). This "train the trainer" model sought to empower departmental analysts to own the training and use it in their departments. Upon completing the Academy, graduates went on to set up training courses for the wider analytical community in their home departments. (The Evaluation Task Force, 2023[29])

### **Identifying and developing human skills for evaluation: Spain's public policy evaluation competence framework**

Spain's 2022 Law on the Institutionalisation of Public Policy Evaluation foresees specific measures to train civil servants in the design, commissioning, implementation and use public policy evaluation. In this context, the Spanish Institute for Public Policy Evaluation (IEPP), in collaboration with the National Institute for Public Administration (INAP) and 11 experts from within the administration, collaboratively developed a public policy evaluation competence framework. The final framework includes a set of competences (e.g. common, specialised and governance competences), job categories (e.g. evaluator, evaluation manager, decision maker, member of an evaluation co-ordination unit); and levels of expertise (basic, intermediate and high). Using this framework, INAP and IEPP intend to develop specific training pathways for each job category, enable self-assessment in skills for evaluation, and facilitate the implementation other upskilling actions across government (Instituto Nacional de Administración Pública and Instituto para la Evaluación de Políticas Públicas, 2024[30]).

### **Developing agreements to access data for impact evaluations in Italy: Visit PNNR-Lavoro**

The "Visit PNNR-Lavoro" initiative, a collaboration between the Italian Ministry of Finance and the Italian Social Security Institute (INPS), promotes robust impact evaluation of the 2021-2026 National Recovery and Resilience Plan (NRRP). It facilitates the use of INPS and Ministry of Finance administrative data for research, enabling evaluations of NRRP measures in itinere and ex post. Anonymized, General Data Protection Regulation-compliant individual-level data is made available to researchers selected through special calls for research proposals. A Steering Committee oversees the programme, identifies relevant data and topics, monitors research progress, and ensures the publication of evaluation results.



## What initiatives have OECD member countries introduced?

### Developing strong data frameworks for evaluation: the Irish Data Sharing and Governance Act and New Zealand's Data and Statistics Act of 2022

The Irish government recently implemented several measures to improve public service data and its use for decision making. In this context, it developed a complete data framework including a dedicated legislation (the Data Sharing and Governance Act), a strategy (the Public Service Data Strategy) and a data governance board and committee. The law and the strategy create clear rules and give direction for institutions on how and when to share their data while increasing citizen transparency. Through the Data Governance Unit in the Office of the Government Chief Information Officer, institutions can be supported in better understanding which data can be accessed for research, particularly in line with the European Union's General Data Protection Regulation (GDPR).

In New Zealand, the Data and Statistics Act also modernises the framework for accessing data for research, and provides safeguards and protections to ensure public trust and confidence in the collection and use of data for official statistics and research. Part 5 of the Act provides rules and guidelines regarding the access to data for research, and specifies that government statisticians may provide access to data for research if the proposed research is in the public interest, and the access is subject to appropriate measure to protect the privacy, confidentiality and security of the data (Government of New Zealand, 2022[31]).



- Adopt evaluation competence frameworks and conduct evaluation skills development activities to help guide the recruitment, training, and professional development of evaluators.
- Deliver training to managers and senior civil servants on the use of evaluations.
- Implement regulation or strategies to facilitate and govern access to data for evaluation.
- Develop strong monitoring systems and procedures to yield data on policy initiative performance.
- Earmark budgets to conduct policy evaluation, as part of budgetary processes.



## Keep in mind

- Evaluation networks and associations can act as important drivers of professionalisation of evaluation communities inside and outside of government.
- Securing financial and human resources specifically for evaluations underscores their importance and enables government institutions to carry out thorough and rigorous evaluations.
- The existence of single access points where evaluators can easily obtain relevant data, can improve data access and literacy and facilitate more comprehensive and data-driven evaluations.



- OECD Digital Government Toolkit. The toolkit is designed to help governments implement the OECD Recommendation on Digital Government Strategies. By comparing good practices across OECD countries, the toolkit can guide decision-makers in using digital technologies and data to encourage innovation, transparency, and efficiency in the public sector (<https://www.oecd.org/governance/digital-government/toolkit/home/>).
- United Nations Evaluation Group (2016[32]), Evaluation Competency Framework. The report presents the evaluation competencies framework for the United Nations System, which can act as a starting point for the development of national / ministry evaluation competencies framework. The framework can be used for a range of purposes including recruitment and hiring, learning and development, promotion and career advancement, and performance assessment. (<https://tinyurl.com/UNEGcompetency>)
- OECD (2017[33]), Skills for a High Performing Civil Service, OECD Public Governance Reviews, OECD Publishing, Paris. The report looks at the capacity and capabilities of civil servants of OECD countries. It explores the skills required to make better policies and regulations, to work effectively with citizens and service users, to commission cost-effective service delivery, and to collaborate with stakeholders in networked settings. The report also suggests approaches for addressing skills gaps through recruitment, development and workforce management. (<http://dx.doi.org/10.1787/9789264280724-en>)

# Pillar 3: Impact of Public Policy Evaluation

## Why does the impact of public policy evaluation matter?

The use of evaluations can help identify what policies work and why, and feed into practical decision making. Without impactful evaluations and other sources of policy evidence, there is an increased risk of decision making based on instinct rather than facts, potentially leading to ineffective policies. Establishing mechanisms to embed evaluation results in decision-making processes, packaging and presenting evidence from evaluations in a user-friendly way and fostering demand for evaluations are key elements to overcoming this challenge.

While most OECD member countries use evaluations to some extent for different purposes (e.g., to inform the development of the government programmes or other strategic plans, to inform budgetary decision making); the use of evaluations – and more broadly of evidence – in policymaking continues to represent a pervasive challenge. This is in part illustrated by the frequent absence in OECD member countries of mechanisms to promote the use of evaluation results in high-level policy discussions.

Evaluations can be of use at many levels of decision making. They can support programme managers to inform operational decision making for instance, as well as to introduce changes to programme design and delivery mechanisms. They can also be used by parliamentarians and policymakers to identify the most fit for purpose policy solutions to a given challenge (i.e. ex ante impact assessment). More broadly, evaluation represents one of the building blocks of a strong performance management and spending system. In combination with other tools such as spending reviews and performance audits, evaluations can support decisions on where and how to allocate financial resources in the most cost-effective manner.



## What the OECD survey on Public Policy Evaluation says about quality in member countries

71% use **evaluation results** to inform budgetary decisions

65% **adopted high-level guidance** framing policy evaluation across government

60% make evaluations **publicly available by default**

40% use **evidence synthesis** methodologies

23% use **systematic follow-up** mechanisms to evaluations

## Dimension 6: Use in Decision Making

### What does the Recommendation say?

Countries should establish institutional mechanisms to embed evaluation in decision-making processes, both at the organisational level and across government.

### What is the aim?

**Improving the timing, type and co-ordination of policy evaluations conducted across government, ensuring they are effectively integrated into decision-making processes.**

There is a need to ensure adequate co-ordination of evaluations to enhance the potential for use in different decision-making contexts (e.g. reform or renewal of the initiative, budgetary planning), the timely production of relevant bodies of evidence, as well as avoiding redundancies.

**Facilitating the use of evaluation evidence in discussions and decisions taking place at the highest levels of government.**

Different mechanisms can facilitate the use of evaluation findings in the context of decisions regarding government-wide policy priorities at the highest political level, such as Councils of Ministers or equivalent bodies. Other measures can enhance the use of evaluation evidence by Parliaments, as they decide on what financial resources to allocate across different government priorities.

**Ensuring ownership of evaluation recommendations and accountability for the implementation of follow-up actions to address them.**

The owners and managers of public policies undergoing evaluations require the right incentives to take ownership of evaluation recommendations, and be made accountable for the execution of actions aimed at addressing them. Without this, evaluations are unlikely to lead to any tangible change in the short, medium or long term.



- Are evaluations being delivered in time for their use in key policy and budgetary decisions, and according to a high-level agenda and planning?
- Are evaluations being used in decision (policy and budgetary)-making processes as a result of explicit governance mechanism allowing or requiring to do so?
- Are decision makers systematically responding to the results of evaluations as a result of evaluation follow-up mechanisms?



## What initiatives have OECD member countries introduced?

### **High level guidance for evaluation across government: the 2022 Periodic**

#### **Evaluations Regulation in the Netherlands**

The Periodic Evaluation Regulations adopted by the Netherlands in 2022 illustrate how governments can provide high-level guidance to carry out evaluations across government through a dedicated regulatory framework. By providing clear definitions and establishing the use of the Strategic Evaluation Agendas (SEA) throughout government, the regulation seeks to ensure comprehensive assessments of policy effectiveness and efficiency. The regulation also intends to foster a culture of continuous improvement and informed decision making by integrating evaluation into various aspects of governance, including subsidy schemes, tax arrangements, legislation, and administrative bodies (Government of the Netherlands, 2022[34]).

### **Cross-government guidance to support the use of evaluation: Canada's Policy on Results**

The Government of Canada's Policy on Results, launched in July 2016, is managed by the Canadian Government's Treasury Board. It sets out the fundamental requirements for Canadian federal departmental accountability for performance information and evaluation, while highlighting the importance of results in management and expenditure decision making, as well as public reporting. Under the Policy, each government department is mandated to establish an evaluation unit as well as a departmental results framework. It also requires evaluations of grants and contribution programmes every five years, and promotes the use of evaluation results in policy decisions by requiring department heads to submit these evaluation along with proposals to the Treasury Board for new spending (Government of Canada, 2016[35]).

### **Strengthening the evaluation – budget planning nexus: Chile's use of evaluation evidence in budgetary and spending processes**

For the past three decades Chile has developed a performance-based budgeting system which draws from the results of evaluations of central government priorities. As part of this "system of evaluation and management control" Chilean authorities are able to link evaluation findings and evidence to decisionmaking processes at the highest level, in a quest to improve the effectiveness of policymaking and management and to make the budget results-oriented.

### **The use of follow-up mechanisms to ensure implementation of evaluation recommendations: the case of Lithuania**

As part of its strategic governance approach, the Lithuanian government adopted a specific mechanism to ensure the follow-up of evaluation recommendations. Through this mechanism, the Ministry of Finance is responsible for monitoring the implementation of evaluation recommendations for all evaluations carried out as part of its Evaluation Plan. In addition, administrative rules linked to investment programmes financed by the European Union's Structural Funds also require State institutions to plan measures for the implementation of evaluation recommendations, which are also monitored by the Ministry of Finance.



- Develop high level guidance providing clear directions on when to conduct policy evaluation and what type of evaluation is needed, in order to adapt the timing of evaluations to feed into decisionmaking processes.
- Adopt mechanisms which incorporate evaluations into decision-making processes (e.g. including evaluation results in policy briefs and proposals, presentations of results at ministerial level)
- Build evaluations into budgetary planning and management processes (e.g. strengthening links between evaluations and spending reviews).
- Implement evaluation recommendation follow-up and monitoring systems.



## Keep in mind

- Formulating clear and specific evaluation questions is key to ensuring relevant and useful findings.
- Understanding stakeholders' needs and concerns ensures that evaluations are relevant and makes findings more likely to be utilised. A specific measure to consider is to surface the questions that senior decision makers seek to answer with evaluation.

- Ensure evaluations are timed to coincide with critical decision-making stages and that evaluation evidence is integrated into ministerial and government workflows (reflecting it in decision documents and briefs).
- Avoid the tendency to commission evaluations in areas where positive results are expected and to avoid more sensitive or underperforming areas (i.e., positivity bias).



- Better Evaluation, Supporting the use of M&E evidence: A summary of the key principles and strategies to support the use of M&E evidence in decision making. (<https://tinyurl.com/betterevalimpact>)
- Budgeting and Public Expenditures in OECD Countries 2019: A comprehensive view of practices and developments in the governance, implementation and performance of budgeting across OECD countries; including the use of evaluation in budgetary processes. (<https://doi.org/10.1787/9789264307957-en>)
- OECD Mobilising Evidence for Good Governance: Taking Stock of Principles and Standards for Policy Design, Implementation and Evaluation: An overview of key governance principles and standards of evidence – including evaluation - to promote evidence-informed policymaking. (<https://doi.org/10.1787/3f6f736b-en>)

## Dimension 7: Accessibility and Dissemination of Findings and Results

### What does the Recommendation say?

Countries should provide easy access to evaluations and present the findings deliberately in order to improve the uptake of evaluation results.

### What is the aim?

**Ensuring that evaluation evidence is easily available and in a format that facilitates and encourages its use in decision-making and planning processes.**

Policymakers looking to use evaluation findings and support the decision-making process through the use of evidence should be able to easily identify, access, and consume them, as well as adapt them to their needs, contexts and purposes. Evaluation reports often remain hidden from potential users and the wider public, and their structure and content can frequently make it challenging to identify and extract relevant information for third parties.

**Ensuring potential users are aware of evaluations, and support them in conducting evidence-informed policymaking.**

The automatic publication of evaluation results should be considered a cornerstone of evaluation systems. Without it, the learning and accountability potential of evaluation is severely hindered. Thus, to guarantee the use and application of evaluation findings, evaluations should be publicly disclosed and known to all relevant stakeholders.

**Consolidating broad bodies of evidence on a single policy topic into simple and easy to use syntheses for policymakers and other users of evidence.**

Evaluation and evidence synthesis methodologies can greatly contribute to enabling the use of evidence in decision making. By collecting, curating, assessing and synthesising existing evidence on a common topic, such approaches are more likely to be consumed and accessed by time and resource-restricted policymakers.

**Evidence synthesis methodologies** are tools to combine and rate evidence from multiple studies that have evaluated the same issue, to come to an overall understanding of what they have found in terms of policy results and to provide an assessment of the robustness of these results.



- Do all potential users, including the general public, have access to evaluation findings through the open publication of all evaluations?
- Is the use of evaluations enabled by user-friendly and accessible results reporting and communication methods?
- Is the use of evaluations in policymaking enabled by the development of evidence synthesis methodologies and their inclusion in existing decision-making tools?



## What initiatives have OECD member countries introduced?

### **Publication and dissemination of federal evaluations: the “Budgetary Transparency” portal of the Mexican Ministry of Finance and Public Credit**

All evaluations of the Ministry of Finance and Public Credit’s budgetary programmes since 2007 are published on the evaluation section of the Ministry’s “Budgetary Transparency” platform. The platform acts as a broader budgetary expenditure observatory of the Ministry, seeking to promote accountability and access to information in open formats to improve the public’s understanding of public spending and budgets. Evaluations can be searched and downloaded by area, programme, and type of evaluation. The platform also contains a database of evaluation recommendations (Presidencia de la Republica and Secretaria de Hacienda y Credito Publico, n.d.[36]).

### **Improving the access and use of evaluation results: the French public policy evaluation platform and Spain’s interactive evaluation dashboard**

The French Supreme Audit Institution (Cour des Comptes) launched a platform in 2024 which collects and centralises more than 900 evaluations with a focus on assessing impact, conducted by 37 French institutions between 2008 and 2023 covering 16 policy areas (Cour des Comptes, 2024[37]). These can be filtered by year, policy area, institution and methodology. The platform collects both qualitative impact evaluations and more experimental evaluations looking at establishing causal relationships. In addition, to support policymakers in accessing the key results, the platform also provides a one-page summary with the key results and the recommendations for each of the evaluations.

Similarly in the Spanish Independent Fiscal Institute in Spain (AIReF) developed an interactive tool known as the Observatory of Evaluations Findings-Proposals (AIReF, 2024[38]), summarising key findings and recommendations from its evaluations. The tool highlights the main proposals, explains them briefly, and identifies the responsible actors for implementation. Each evaluation has a summary page with essential information on policy costs and the evaluation methodology. This tool facilitates quick access to key results and ensures transparency of evaluation findings.

### **Synthesising and improving access to evidence for decision making: the What Works Clearinghouse (WWC) of the Department of Education in the United States**

The WWC acts as a source of scientific evidence on education programmes, products, practices, and policies; designed to support evidence-informed policymaking on education. It is responsible for reviewing available research, determine which studies meet rigorous standards, and summarise the findings. It conducts systematic reviews of bodies of evidence, as well as reviews of individual studies (Institute of Education Sciences, n.d.[39])



Act

- Build evaluation repositories, clearinghouses and other evaluation knowledge infrastructure.
- Systematically publish evaluation results and develop effective communication strategies around them, including targeted dissemination plans.
- Conduct evidence gaps analyses and evidence syntheses.
- Adopt user-friendly, short, and visually impactful evaluation results presentations.



## Keep in mind

- Investing in communication skills within evaluation teams as well as engaging with communication experts can contribute to clearly and compellingly conveying results to various stakeholders.
- Storytelling and visualisation can make complex findings more accessible and engaging.
- Unified portals and metadata systems for accessing evaluation results may facilitate easier retrieval and use of evaluation findings.



Learn

- Better Evaluation, Develop reporting media: Guidance and principles on the different types of evaluation reports which can be developed for different audiences. (<https://tinyurl.com/betterevalmedia>)
- Fresh Spectrum Data Visualisation Blog and newsletter: A variety of resources to learn and improve skills and gain insights on how to visualise and present the results of evaluations. (<https://freshspectrum.com/>)
- Kojoyan, S. et al. (2021[40]), OED Guidance Note on Evaluation Synthesis. FAO Office of Evaluation: A guidance note providing a reference for evaluators with respect to planning, formulating and conducting evaluation syntheses. (<https://tinyurl.com/KojoyanFAO>)
- Saran, A. and H. White (2018[41]), "Evidence and gap maps: a comparison of different approaches": A paper reviewing the range of evidence mapping methodologies adopted across different sectors, and developing general principles and standards for maps. (<http://dx.doi.org/10.4073/cmdp.2018.2>)

# Conclusion & additional support

This toolkit offers policymakers operational and practical guidance on how to implement the principles of the OECD Recommendation on Public Policy Evaluation (see Annex A for the main text of the Recommendation). The Recommendation seeks to strengthen the practices and institutions aimed at driving the use of public policy evaluation throughout government and beyond, to support evidenceinformed policymaking.

The toolkit contains a detailed explanation of the different principles of the Recommendation, as well as the key concepts behind them. In addition, it offers practical tips, references to learning material, as well as examples of real-life cases illustrating how these principles are put into practice by different OECD Member countries. Finally, the different Pillar chapters contain a simple set of self-assessment questions for users seeking to assess their situation regarding the principles contained in the Recommendation. However, for users seeking to conduct a more in-depth assessment of their public policy evaluation capacities from an institutionalisation (i.e. Pillar 1), quality (i.e. Pillar 2), or impact (i.e. Pillar 3) perspectives, Annex B contains a dedicated and detailed self-assessment tool. Further information on how the tool can be used and for what purposes is also found in Annex B.

The remaining Annexes of the toolkit provide additional information on the evaluation guidelines and the national evaluation champions in different OECD Member countries.

Public policy evaluation can be more impactful when deployed alongside other performance and budget management solutions such as performance budgeting, spending reviews and policy monitoring. The alignment of these different solutions can bring about many benefits including the development of virtuous cycle of evidence-based policymaking, informed budgetary decision making and fiscally responsible governance. Additional knowledge developed by the OECD in these fields can be found at:

- Public policy monitoring and evaluation: <https://www.oecd.org/en/topics/public-policy-monitoringand-evaluation.html>
- Performance budgeting: <https://www.oecd.org/en/topics/performance-budgeting.html>
- Spending reviews: <https://www.oecd.org/en/topics/spending-reviews.html>

## The OECD Recommendation on Public Policy Evaluation: Key OECD Member facts and good practices



Recommendation Pillars



Figures and facts on OECD member countries\*



Good practices

# References

- AIReF (2024), Observatory of Findings-Proposals Evaluations, [38]  
<https://www.airef.es/es/buscadorhallazgos/>.
- Assemblée Nationale (2024), L'évaluation des politiques publiques, [7]  
<https://www.assembleenationale.fr/dyn/synthese/fonctionnement-assemblee-nationale/evaluation-politiquespubliques-controle-gouvernement/l-evaluation-des-politiques-publiques>.
- Aston, T. and M. Apgar (2023), United Kingdom Evaluation Society Training on Evidence Standards. [42]
- Australian Centre for Evaluation (2023), Commonwealth Evaluation Toolkit, [21]  
<https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit>.
- Commission des finances, de l'économie générale (2023), L'évaluation des politiques publiques 2023. [8]
- Congress (2018), Foundations for Evidence-Based Policymaking Act of 2018 Public Law No: 115-435 (01/14/2019), <https://www.congress.gov/bill/115th-congress/house-bill/4174>. [22]
- Cour des Comptes (2024), Plateforme des évaluations, [37]  
<https://www.ccomptes.fr/fr/plateformescitoyennes/plateforme-evaluations-politique-publique>.
- European Commission (n.d.), Regulatory Scrutiny Board, [24]  
[https://commission.europa.eu/law/lawmaking-process/regulatory-scrutiny-board\\_en](https://commission.europa.eu/law/lawmaking-process/regulatory-scrutiny-board_en)  
(accessed on 23 October 2024).
- Evaluation Task Force (2024), Evaluation Task Force, <https://www.gov.uk/guidance/what-works-network>. [11]
- Government of Canada (2016), Policy on Results, <https://www.tbs-sct.canada.ca/pol/doceng.aspx?id=31300> (accessed on 14 May 2024). [35]
- Government of Japan (2001), "Government Policy Evaluations Act (Act No. 86 of 2001)", [http://Government Policy Evaluations Act \(Act No. 86 of 2001\)](http://Government Policy Evaluations Act (Act No. 86 of 2001) (accessed on 13 May 2024)) (accessed on 13 May 2024). [3]
- Government of New Zealand (2022), Data and Statistics Act 2022. [31] Government of the Netherlands (2022), Regeling periodiek evaluatieonderzoek 2022, <https://wetten.overheid.nl/BWBR0046970/2022-07-27> (accessed on 14 May 2024). [34]
- Head of State (2022), Law 27/2022, of December 20, 2022, on the institutionalization of the evaluation of public policies in the General State Administration, <https://www.boe.es/eli/es/l/2022/12/20/27/con>. [4]

- Institute of Education Sciences, U. (n.d.), What Works Clearinghouse, [http://Institute of Education Sciences \(IES\) within the U.S. Department of Education](http://Institute of Education Sciences (IES) within the U.S. Department of Education) (accessed on 14 May 2024). [39]
- Instituto Nacional de Administración Pública and Instituto para la Evaluación de Políticas Públicas (2024), “Marco de Competencias en Evaluación de Políticas Públicas”, <https://www.inap.es/documents/10136/1976576/Marco+de+evaluación/d0c4271c-ce2f-e5d5-9bcd-60cac40fa8bd> (accessed on 9 July 2024). [30]
- INTOSAI (2019), GUID 9020: Evaluation of Public Policies, <https://www.issai.org/wpcontent/uploads/2019/08/GUID-9020-Evaluation-of-Public-Policies.pdf>. [12]
- Kojoyan, S. et al. (2021), , <https://www.fao.org/3/cb7077en/cb7077en.pdf>. [40]
- Lázaro, B. (2015), Comparative study on the institutionalisation of evaluation in Europe and Latin America, EUROsociAL Programme, [https://sia.eurosocial-ii.eu/files/docs/1456851768-E\\_15\\_ENfin.pdf](https://sia.eurosocial-ii.eu/files/docs/1456851768-E_15_ENfin.pdf). [5]
- Ministry of Economic and Finance (2023), Criteria and methodologies for the definition of processes and activities for the analysis and evaluation of expenditure, <https://area.rgs.mef.gov.it/canali/498/bacheca>. [15]
- Ministry of National Planning and Economic Policy (2019), Participatory Evaluation Guide, [https://evalparticipativa.net/wpcontent/uploads/2019/06/26.guia\\_de\\_evaluacion\\_con\\_participacion.pdf](https://evalparticipativa.net/wpcontent/uploads/2019/06/26.guia_de_evaluacion_con_participacion.pdf). [16]
- Nutley, S., A. Powell and H. Davies (2013), What counts as good evidence? Alliance for Useful Evidence, Research Unit for Research Utilisation (RURU), School of Management, University of St Andrews, <https://media.nesta.org.uk/documents/What-Counts-as-Good-EvidenceWEB.pdf>. [25]
- OECD (2022), Recommendation of the Council on Public Policy Evaluation, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0478> (accessed on 5 June 2024). [2]
- OECD (2021), Applying Evaluation Criteria Thoughtfully, OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>. [19]
- OECD (2021), OECD Regulatory Policy Outlook 2021, OECD Publishing, Paris, <https://doi.org/10.1787/38b0fdb1-en>. [14]
- OECD (2020), Improving Governance with Policy Evaluation: Lessons From Country Experiences, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/89b1577d-en>. [6]
- OECD (2020), Policy Framework on Sound Public Governance: Baseline Features of Governments that Work Well, OECD Publishing, Paris, <https://doi.org/10.1787/c03e01b3-en>. [1]
- OECD (2020), Reviewing the Stock of Regulation, OECD Best Practice Principles for Regulatory Policy, OECD Publishing, Paris, <https://doi.org/10.1787/1a8f33bc-en>. [18]

- OECD (2017), Skills for a High Performing Civil Service, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/9789264280724-en>. [33]
- OECD (2010), Quality Standards for Development Evaluation, DAC Guidelines and Reference Series, OECD Publishing, Paris, <https://doi.org/10.1787/9789264083905-en> [26]
- Office of Management and Budget (2020), Phase 4 Implementation of the Foundations for Evidence-Based Policymaking, <https://www.whitehouse.gov/wp-content/uploads/2020/03/M20-12.pdf>. [23]
- Patton, M. (2002), Utilization-focused evaluation checklist, <https://wmich.edu/evaluation/checklists>. [20]
- Presidencia de la Republica and Secretaria de Hacienda y Credito Publico (n.d.), Portal de Transparencia Presupuestaria, <https://www.transparenciapresupuestaria.gob.mx/> (accessed on 14 May 2024). [36]
- Preskill, H. and N. Jones (2009), A Practical Guide for Engaging Stakeholders in Developing Evaluation Questions, Robert Wood Johnson Foundation, [https://www.betterevaluation.org/sites/default/files/Preskill\\_A\\_Practical\\_Guide.pdf](https://www.betterevaluation.org/sites/default/files/Preskill_A_Practical_Guide.pdf). [17]
- Revillard, A. (2023), Policy Evaluation: Methods and Approaches, <https://www.sciencespo.fr/liepp/fr/node/5899.html>. [27]
- Saran, A. and H. White (2018), "Evidence and gap maps: a comparison of different approaches", Campbell Systematic Reviews, Vol. 14/1, pp. 1-38, <https://doi.org/10.4073/cmdp.2018.2>. [41]
- Speer, S. (2015), "The Growing Role of Evaluation in Parliaments: Holding Governments Accountable", International Review of Administrative Sciences, Vol. 81/1, pp. 37-57. [13]
- Swedish National Audit Office (2024), Sweden's international development aid - monitoring, evaluation and reporting, [https://www.riksrevisionen.se/download/18.428f0ec018eab1b53857ba85/1712565464289/RiR\\_2024\\_4\\_summary.pdf](https://www.riksrevisionen.se/download/18.428f0ec018eab1b53857ba85/1712565464289/RiR_2024_4_summary.pdf). [10]
- Swedish National Audit Office (2022), On shaky ground - decision support ahead of major reforms, [https://www.riksrevisionen.se/download/18.1ef1d69b1817fe51b21303d8/1655822160799/RiR\\_2022\\_15\\_summary.pdf](https://www.riksrevisionen.se/download/18.1ef1d69b1817fe51b21303d8/1655822160799/RiR_2022_15_summary.pdf). [9]
- The Evaluation Task Force (2023), The Evaluation Task Force's (ETF) Evaluation Academy introduces a new way to train evaluators in government. [29]
- UNEG (2016), Evaluation Competency Framework, United Nations Evaluation Group, <https://www.unevaluation.org/document/detail/1915>. [32]
- Varazzani, C. et al. (2023), "Seven routes to experimentation in policymaking: A guide to applied behavioural science methods", OECD Working Papers on Public Governance, No. 64, OECD Publishing, Paris, <https://doi.org/10.1787/918b6a04-en>. [28]

# Annex A Main text of the OECD Recommendation on Public Policy Evaluation (OECD/LEGAL/0478)

*The following text presents the main elements contained in the Recommendation of the Council on Public Policy Evaluation. The full text of the Recommendation in English and French, and the unofficial translations to other languages can be found at <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0478>.*

The Council, on the proposal of the Public Governance Committee and Committee of Senior Budget Officials:

**I. RECOMMENDS** that Members and non-Members having adhered to this Recommendation (hereafter the “Adherents”) institutionalise public policy evaluation from a whole of government perspective. To this effect, Adherents should:

**a.** Conduct and use evaluations across government ensuring that they are carried-out in a systematic manner and that their results are used in policy and budgetary decision making. In particular, Adherents should:

**i.** Designate evaluation champions to co-ordinate evaluations across institutions and advise on best practices to promote their quality and use.

**ii.** Define and assign institutional responsibilities for conducting policy evaluations.

**b.** Foster a culture of learning and accountability by promoting demand for, and ownership of, evaluations within and beyond the executive. In particular, Adherents should:

**i.** Promote the role of both internal and external, national and international, knowledge brokers to strengthen the relationship between evidence from evaluations and its users, including citizens.

**ii.** Offer opportunities for the legislative body to review and discuss policy evaluations.

**II. RECOMMENDS** that Adherents promote the quality of public policy evaluations. To this end, Adherents should:

**a.** Actively plan, design and manage evaluations so that they are timely and proportionate to the intended objectives, taking into account the needs of the primary users and the types of intended uses, and ensuring that results can be trusted by stakeholders. In particular, Adherents should:

**i.** Plan evaluations early by building provisions for evaluations into public interventions from the start, in order to improve their design, collect data on their implementation, and ensure that evaluation results are robust and available in a timely fashion.

ii. Design and implement evaluations that are proportionate and appropriate for the likely use, by adapting the aim, scope and analysis of the evaluation, its format and resources, to the needs of its primary users and the types of intended uses.

iii. Engage relevant stakeholders in the evaluation processes from the outset in order to create ownership for change and trust in evaluation results.

b. Establish quality standards and mechanisms for evaluations to generate robust and credible evaluation results that can be trusted and used with confidence. In particular, Adherents should:

i. Develop guidelines to ensure that evaluation designs, data collection processes and analytical methods, adhere to methodological best practices.

ii. Adopt professional and ethical standards for evaluators to ensure that they meet high criteria for integrity and independence, as well as for knowledge of evaluative methods and culturally appropriate approaches, and that they safeguard the dignity, rights, safety and privacy of participants and other stakeholders when they conduct evaluations.

iii. Promote the functional autonomy of evaluations, by safeguarding the autonomy of external evaluations through oversight of the commissioning and evaluation processes, and by providing internal evaluations team with a high degree of autonomy in the use of available resources and in deciding what studies to conduct and how.

iv. Ensure that evaluations are able to withstand external scrutiny, such as through peer review, and that they can be assessed against pre-defined quality criteria.

c. Develop institutional skills and capacities to conduct, commission and use evaluations effectively and in a credible manner. In particular, Adherents should:

i. Build public sector skills for evaluation by conducting regular training, recruiting and retaining employees with the adequate skills or collaborating with academia, the private sector and other jurisdictions to improve the availability of these skills.

ii. Ensure the availability of high quality, timely, accessible, disaggregated and reusable results, performance and administrative data for policy evaluation.

iii. Provide institutions with appropriate resources to manage, carry-out and use policy evaluations.

**III. RECOMMENDS** that Adherents conduct public policy evaluations that impact decision making. To this effect, Adherents should:

**a.** Establish institutional mechanisms to embed evaluation in decision-making processes, both at the organisational level and across government. In particular, Adherents should:

**i.** Provide high-level guidance, such as in a legal or policy framework or in a multiannual evaluation agenda, on when to conduct policy evaluation and what type of evaluation is needed, in order to adapt the timing of evaluations to feed into decision making processes, focus the analysis where it is most needed, coordinate efforts for cross-sectorial evaluations, and avoid overlaps.

**ii.** Incorporate the use of evaluation results into decision making including through the policymaking and budgetary processes.

**iii.** Establish follow-up mechanisms for decision makers to respond to the results of evaluations, by defining a course of action where relevant, and assigning responsibilities for implementing and tracking recommendations.

**b.** Provide easy access to evaluations and present the findings deliberately in order to improve the uptake of evaluation results. In particular, Adherents should:

**i.** Make the result of evaluation findings and recommendations public by default.

**ii.** Tailor the way evaluation evidence is presented and communicated to its potential users, in terms of timing, communication channel, format and messaging, by developing a dissemination strategy.

**iii.** Make use of evidence synthesis methodologies to aggregate evaluation findings and assess them in a systematic manner.

# Annex B Self-assessment tool

## What is the self-assessment tool?

The assessment element of this Toolkit is one its key features. It is designed for users to take stock of the current strengths and weaknesses of their evaluation practices and identify areas/means of improvement. The toolkit offers two assessment tools:

- A simple set of assessment questions are presented under each of the dimensions covered by the main body of the Toolkit. These questions can be used as primers for an initial reflection on each of the Recommendation's dimensions.
- In addition, a detailed assessment is presented below. This tool allows users to perform an assessment of their situation and level of maturity against each of the Recommendation's criteria, using a set of 20 rubrics. The rubrics are organised according to the Recommendations three pillars and seven dimensions.

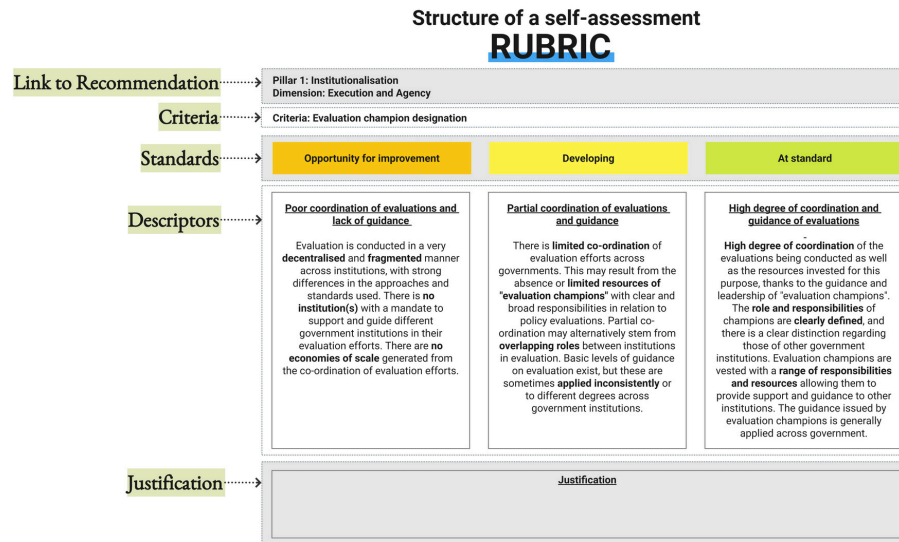
## What are self-assessment rubrics?

Rubrics are a form of qualitative scale that denotes a level of performance. They explain what each criteria means and make clear the reasoning behind an assessment. Each rubric included in the self-assessment tool includes several components (see Figure below):

- The **link to the Recommendation** specifies how the criteria being assessed is tied to the general structure of the Recommendation on Public Policy Evaluation. Specifically, this indicates which of the Recommendation's three Pillars and Seven dimensions the criteria being assessed is under.
- **Criteria** are the aspects of performance that are of interest to the self-assessment. They represent the basic unit of assessment of the tool. Individual rubrics can be used assess a total of 20 different criteria linked to the Recommendation on Public Policy Evaluation. The aggregated results of criteria assessment can allow to assess the recommendations dimensions (7) or Pillars (3) at a more general level.
- **Standards** express the levels of performance or quality for each criterion, based on the results of the assessment and the evidence gathered. For each rubric / criterion, users should select one of three standards:
  - **Opportunity for improvement:** indicates the existence important gaps between the existing situation and / or capacities, and the standards set by the Recommendation.
  - **Developing:** indicates an intermediate level of achievement of the standards set by the Recommendation, with room for the introduction of additional improvements.
  - **At standard:** indicates that the country or institution is operating at the standard set by the Recommendation.
- **Descriptors** are examples or descriptions of what each standard looks like for each criterion in the rubric. These should allow users to select the most adequate standard for their appraisal of each criterion.

- **Justifications** should provide an explanation for the selected standard, explaining how the conclusion was reached and presenting any potential nuance or exception to the assessment. Justifications should systematically be provided and document for each of the criteria covered by the assessment.

As opposed to other assessment methods or tools (e.g., checklists or scales), rubrics are a more appropriate way of assessing areas where boundaries not clearly delineated. Rubrics are also more culturally flexible and adaptable (Aston and Apgar, 2023[28])



## What types of evaluations does the tool apply to?

The definition given to evaluation in the context of this toolkit is purposely broad, and meant to cover a broad range of evaluation approaches from a methodological (e.g. theory-based, experimental), policy cycle perspective (e.g. ex-ante, ex-post), and instrument perspective (e.g. policies, programmes, legislation). As part of any self-assessment exercise, users should reflect on the types of evaluations they wish to apply the toolkit to and adapt its use accordingly. The OECD has developed additional guidance focusing specifically on regulatory impact assessment which may be more relevant and tailored to the assessment of capacities in this specific area of evaluation. Users may rely on the justifications section of the assessment rubrics to describe any differences or nuances to be given to the assessment of different types of evaluations within their governments or institutions.

## How can this self-assessment tool be implemented?

There is no standard approach to the use of the self-assessment tool. Some users may use it to organise an internal discussion on existing capacities, while others may use it as the basis to conduct a detailed and full-fledged analysis of their evaluation systems. As such, the scope and depth of the self-assessment exercise may vary and can be adjusted based on the context of the user and the intended use of the results. However, the following steps and principles may facilitate its effective use and implementation.

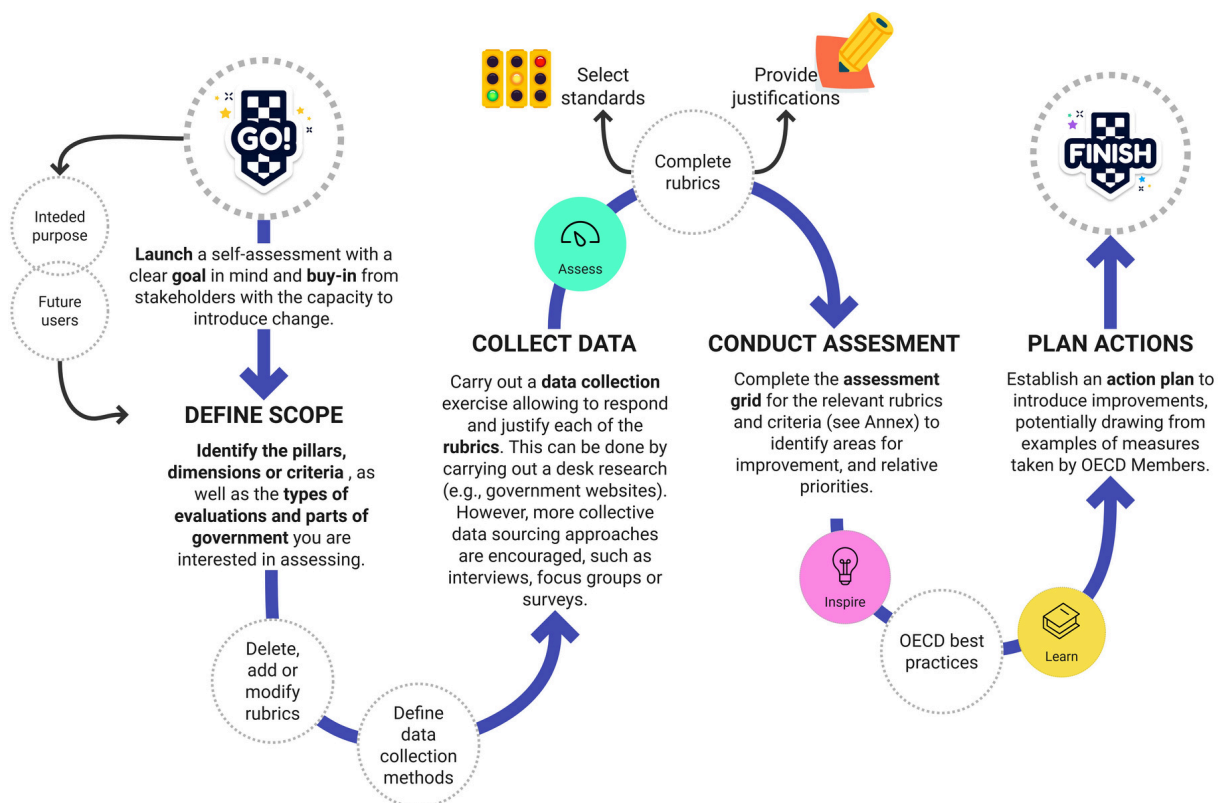
- Governments and / or government institutions may wish to undertake a self-assessment exercise in the context of broader efforts to enhance their public policy capacities. The undertaking of such an exercise should be accompanied by the intention to introduce improvements to existing evaluation systems, an accordingly, support and buy-in from agents who are in position to introduce such improvements.

- While some users may conduct the exercise using internal resources, others may seek the assistance of a third party (e.g. researcher, consultant) when conducting the self-assessment exercise. This decision will depend largely on the nature, the scope and depth of the exercise being conducted, as well as the resources available to do so. A robust yet-straightforward self- assessment exercise can be performed on the basis of limited resources (internal or external) and should not systematically require an extensive investment on behalf of the user.
- The self-assessment tool and its individual rubrics should imperatively be tailored and adapted to the context within which it will be applied, as well as to the purpose of the exercise. In this process, some rubrics may be considered inadequate, or may require adjusting. Additional rubrics may be added. The number of standards, as well as the descriptors used can be modified or adjusted.
- Scoping the self-assessment exercise to be conducted is a key step to ensure an adequate outcome. As part of the self-assessment design phases, users clearly define the scope of the self- assessment they wish to perform, and specifically, identify the Recommendation Pillars and / dimensions they are interested in assessing. Alternatively, an extensive assessment can be conducted for all Pillars / dimensions. Further, the types of evaluations as well as the parts of governments to be covered by the assessment should be clearly identified.
- Once the scope is defined, different data collection methods such as interviews, focus groups, or surveys can be conducted to assess the different rubrics under each of the selected Pillars / dimensions. The specific data collection tools and data sources (e.g. primary secondary) and the stakeholders to be mobilised to perform the assessment should be identified on a case-by-case basis, in line with the objectives of the exercise and its end-use and users. A simple self-assessment exercise can be conducted drawing from the experience and views of a team of evaluation specialists in a single government institution; while a more comprehensive exercise can involve a broader set of stakeholders such as policy officers, programme managers, evaluations specialists from different government agencies and institutions.
- The data collected should enable the selection of a standard for each rubric (i.e., opportunity for improvement, developing, at standard) which best qualifies the user's situation. In addition, a justification should also be drafted for this choice, which can also provide necessary nuances or exceptions.
- The assessment grid (see page 56) can be used to collect and centralise the standards selected for each rubric, providing an overview of the relative strengths and weaknesses identified. This can be done by simply checking the selected standard for each criteria / rubric.

- At the end, an action plan can be developed to address priority areas or low hanging fruit for improvement. Inspiration on the actions to be taken can be drawn from the examples presented in the Toolkit.

While the self-assessment provides an objective exercise to determine a country or institution’s level of capacity or maturity with regards to the different principles and dimensions of the Recommendation, its use should nevertheless always be tailored to the local context its being used in. In addition, the self- assessment does not mean to promote the introduction of specific formal solutions when it comes to enhancing public policy evaluations. Instead, it’s ambition is to help countries and institutions focus on a set of desirable outcomes such as institutionalisation, co-ordination, sustainability, and inclusiveness when thinking about how to further strengthen their capacities.

### Suggested self-assessment implementation roadmap



# Rubrics

## Pillar 1 Dimension 1: Execution and Agency

### Evaluation champion designation

● Opportunity for improvement

● Developing

● At standard

#### Poor co-ordination of evaluations and lack of guidance

Evaluation is conducted in a very decentralised and fragmented manner across institutions, with strong differences in the approaches and standards used. There is no institution(s) with a mandate to support and guide different government institutions in their evaluation efforts. There are no economies of scale generated from the coordination of evaluation efforts.

#### Partial co-ordination of evaluations and guidance

There is limited co-ordination of evaluation efforts across governments. This may result from the absence or limited resources of "evaluation champions" with clear and broad responsibilities in relation to policy evaluations. Partial co-ordination may alternatively stem from overlapping roles between institutions in evaluation. Basic levels of guidance on evaluation exist, but these are sometimes applied inconsistently or to different degrees across government institutions.

#### High degree of co-ordination and guidance of evaluations

High degree of co-ordination of the evaluations being conducted as well as the resources invested for this purpose, thanks to the guidance and leadership of "evaluation champions". The role and responsibilities of champions are clearly defined, and there is a clear distinction regarding those of other government institutions. Evaluation champions are vested with a range of responsibilities and resources allowing them to provide support and guidance to other institutions. The guidance issued by evaluation champions is generally applied across government.

Comments, justification, observations

### Institutional responsibility definition

● Opportunity for improvement

● Developing

● At standard

#### Institutional responsibilities for evaluation are not clearly defined or assigned across government institutions, and evaluation is mostly conducted by one type of actor

There is no formal definition or attribution of institutional evaluation responsibilities across government institutions. Existing evaluation efforts are scarce and fragmented and vary in their quality and the standards to conduct them. There is no possibility of knowing centrally which policies have been evaluated and for what purpose, but it's likely that most policies implemented nationally do not undergo an evaluation.

#### Institutional responsibilities for evaluation are partially defined, and there is a limited diversity of actors involved in evaluation

Some responsibilities for evaluation are defined centrally, in a legal or policy framework or elsewhere. However, certain policy areas have more clearly defined responsibilities than others, leading to varying degrees of use and implementation of evaluations within government. There may be overlaps in the responsibilities of central evaluation champions and those of evaluation units in line ministries, leading to some inefficiencies or duplication of efforts.


#### Institutional responsibilities for evaluation are well defined and attributed, and there is a diversity of actors involved in evaluation

Responsibilities to co-ordinate, commission, conduct and use evaluations are well defined and assigned to institutions across government. These may be included in a cross-governmental legal or policy framework, and offer a complete picture of who does what avoiding potential duplication of responsibilities or unattributed roles; and ensuring the minimum set of evaluation functions to be performed across government.


Comments, justification, observations

**Pillar 1** Dimension 2: Learning and Accountability Culture

**Knowledge broker promotion**

 Opportunity for improvement

 Developing

 At standard

**The supply and demand for evaluation results are not coordinated and there are very few knowledge brokers to carry out this function**

Knowledge brokers do not exist, or are not working in the realm of evaluation. There are no actors performing such a brokerage function between evaluation results and users. There are many missed opportunities for existing evaluations to support key decision-making processes.

**The supply and demand for evaluation results are partially co-ordinated by a limited number of knowledge brokers in some policy areas**


Existing knowledge brokers are concentrated in a limited number of policy areas. These conduct some activities such as organising knowledge dissemination events, but could expand their work into other more complex work (e.g. evaluation syntheses). Their work facilitates the use of evaluation evidence in some key decision-making processes. Outreach efforts focus primarily on policymakers.

**The supply of and demand for evaluation results and evidence are adequately co-ordinated by a community of knowledge brokers**


There is a strong network of knowledge brokers working in different policy areas, providing a range of activities such as training on the use of evaluation to policymakers and synthesising evaluation results. This supports the dissemination of evaluation results and facilitates its use by decision makers in most key decision-making processes, as well as the dissemination of evaluation results among citizens.

**Comments, justification, observations**

**Legislative review opportunity**

 Opportunity for improvement

 Developing

 At standard

**The legislative body only reviews and discusses evaluation results in limited cases**

The legislative body's engagement in policy evaluation is limited, with minimal or no responsibilities related to evaluation processes. There may be no provisions for requiring government evaluations, requesting presentations of evaluation findings, or discussing evaluation results within the legislative body. Absence of evaluation responsibilities and opportunities to discuss evaluation evidence deprives the legislative body of opportunities to leverage evidence-based policymaking, potentially leading to suboptimal policy decisions or missed opportunities for improvement.

**The legislative body reviews and discusses evaluation evidence in some policy areas or for selected legislative projects**




The legislative body engages to some extent in policy evaluation by exercising some responsibilities related to evaluation processes. This may include the ability to require government evaluations, request presentations of evaluation findings, or engage in discussions on evaluation results. However, there may be a small or no internal body dedicated to conducting policy evaluations, and the authority or resources to commission evaluations to external evaluators may be lacking.

**The legislative body frequently reviews and discusses evaluation evidence**




The legislative body actively engages with policy evaluations by exercising multiple responsibilities related to evaluation processes. This includes the ability to require government evaluations, request presentations of evaluation findings, engage in discussions on evaluation results, and potentially have an internal body dedicated to conducting policy evaluations. The legislative body may also have the authority to commission evaluations to external evaluators. Opportunities for legislative review and discussion of evaluation results facilitate evidence-based policymaking, enabling lawmakers to make informed decisions.

**Comments, justification, observations**

**Early evaluation planning and provision**




 Opportunity for improvement	 Developing	 At standard
<p><b>Evaluations are not planned ahead of time and / or as part of policy design, and planning systems and evaluation requirements are absent</b></p> <p>Evaluations are conducted on an ad-hoc basis, and not according to general planning systems. Decisions to carry out evaluations often take place during or after the implementation of policies. There are very few or no evaluation requirements in place. Relevant evidence for decision making is often not available or is produced after the decision-making process takes place.</p>	<p><b>Some evaluations are planned in advance through partial use of planning systems and / or evaluation requirements, or the integration of evaluation in policy design</b></p> <p>Evaluation planning processes are only used in some policy fields or by some line ministries, meaning that only some evaluations are conducted in anticipation of specific evidence needs. Basic evaluation requirements may exist (e.g. policies that result from international commitments), but only apply to a small fraction of policies implemented or financed. Evaluation evidence is only available in anticipation of some policymaking decisions.</p>	<p><b>Evaluations are frequently planned in advance through the combined use of planning systems and evaluation requirements, and the integration of evaluation in policy design</b></p> <p>There is a clear system or processes to ensure that evaluations are planned on time and in parallel to policy design, considering future uses. Requirements on which policies should be evaluated and when are clearly formulated at the central level, allowing policy managers to anticipate the need to conduct evaluations and collect necessary data. Evaluation planning processes are frequent in line ministries and other institutions conducting evaluations.</p>
<p><b>Comments, justification, observations</b></p>		

**Evaluation design relevance**

 Opportunity for improvement	 Developing	 At standard
<p><b>Evaluations are rarely proportionate and appropriate for their likely use, as consideration of use and users in evaluation design is limited</b></p> <p>Evaluations are not adapted or designed with the needs of primary users in mind. The necessary mechanisms and tools to do so do not exist. Evaluations are often performed using a standard approach and method. As a result, their scope / methodology is frequently too large or limited, do not address relevant questions, and / or do not respond to the needs of their future users.</p>	<p><b>Evaluations are sometimes proportionate and appropriate for their likely use, but consideration of use and users in evaluation design is not systematic</b></p> <p>Some evaluations are adapted or designed with the needs of primary users and intended purposes in mind. Mechanisms and tools which allow to do so primarily exist in some line ministries or policy areas or are not used systematically. Primary users and intended uses of evaluations are not systematically identified prior to the launch of evaluations. Tailoring of evaluation design based on future uses and users exists but is not fully widespread.</p>	<p><b>Evaluations are frequently proportionate and appropriate for their likely use given the systematic consideration of use and users in their design</b></p> <p>There are mechanisms and tools in place and in use by evaluation commissioners to consider the potential intended uses and users of evaluations in their design. Uses and users of evaluations are clearly defined before the launch of evaluations. Future users are involved in the design process and / or drafting of the terms of reference. Evaluation methods, scope and resources vary depending on the intended users and uses of the evaluation.</p>
<p><b>Comments, justification, observations</b></p>		




**Pillar 2 Dimension 3: Design, Planning and Management**

**Stakeholder engagement (i.e. direct and indirect beneficiaries, programme owners, evaluation users)**

 Opportunity for improvement	 Developing	 At standard
<p><b>Ownership for and trust in evaluations in evaluations is limited, as relevant stakeholders are rarely engaged in evaluation processes</b></p> <p>Stakeholder consultation and engagement are sporadic or limited, occurring only for select evaluations or in response to specific requirements or circumstances. Engagement activities may be informal, ad-hoc, or reactive, lacking systematic processes or frameworks for stakeholder involvement. While efforts may be made to gather stakeholder feedback, the integration of stakeholder inputs into evaluation processes is inconsistent or minimal.</p>	<p><b>Ownership for and trust in evaluations in strong in some cases, as stakeholder engagement in evaluation processes varies</b></p> <p>Stakeholder consultation and engagement occur for most evaluations, although the extent and depth of engagement, and the types of stakeholders engaged, may vary. While some evaluations involve stakeholders from the outset, others may engage stakeholders at specific stages or on an ad-hoc basis. There are limited systematic procedures or rules in place for this purpose. Efforts are made to incorporate stakeholder feedback and inputs into evaluation processes, but the level of integration may vary depending on the evaluation context.</p>	<p><b>Ownership for and trust in evaluations is strong given the systematic engagement of relevant stakeholders in evaluation processes</b></p> <p>Stakeholder consultation and engagement are integral components of the evaluation process, initiated from the outset of evaluations when relevant. Stakeholders representing diverse perspectives and interests are actively involved in key stages of the evaluation, including design, data collection, analysis, and dissemination; based on clear rules and guidelines. There is evidence of systematic efforts to solicit feedback, address concerns, and incorporate stakeholder inputs into evaluation findings and recommendations.</p>
<p><b>Comments, justification, observations</b></p>		


**Pillar 2 Dimension 4: Quality & Standards**

**Evaluation guideline development**


 Opportunity for improvement	 Developing	 At standard
<p><b>Evaluations are generally designed and conducted using outdated or weak methodological approaches</b></p> <p>There is no formal documentation or framework outlining methodological principles or standards for evaluations at the central level. Evaluation practitioners rely on individual expertise, experience, or ad-hoc guidance from external sources, such as academic literature or professional networks, to inform their methodological approaches. The absence of comprehensive centralised guidelines may result in variability and inconsistency in evaluation practices, with practitioners adopting disparate methods and standards based on personal preferences or institutional norms.</p>	<p><b>The use of methodological best practices for evaluation design and delivery is variable</b></p> <p>Central guidelines offer general guidance on evaluation principles and processes but may not provide detailed instructions or recommendations on specific methodological aspects. While certain methodological considerations may be addressed in the guidelines, coverage may be limited or superficial, leaving gaps in critical areas such as study design or data analysis. The guidelines may emphasise broad principles of evaluation quality and integrity without delving into the nuances of methodological application or implementation. Alternatively, guidelines exist but are not consistently used in practice across different government institutions.</p>	<p><b>Evaluations are very frequently designed and conducted according to methodological best practices</b></p> <p>Flexible but comprehensive central guidelines are developed and accessible to evaluation practitioners, outlining best practices for designing, conducting, and analysing evaluations. The guidelines cover a wide range of methodological aspects, including but not limited to study design, sampling techniques, data collection methods, data analysis approaches, and interpretation of findings. Specific recommendations or standards are provided for ensuring rigor, validity, reliability, and ethical considerations in evaluation methodologies.</p>
<p><b>Comments, justification, observations</b></p>		

**Pillar 2** Dimension 4: Quality & Standards

**Professional standards adoption**

 Opportunity for improvement

 Developing

 At standard

**Evaluation integrity and independence are not guaranteed given the absence of ethical and / or professional standards for evaluators**

There are no specific professional standards or accreditations required for government evaluators, and ethical standards for civil servants may be lacking or not formally integrated into evaluation practices. Evaluations rely instead on ad-hoc practices or informal norms. This increases the risk of substandard evaluation practices, including biases, conflicts of interest, and ethical violations, jeopardising the credibility and reliability of evaluation findings.

**Evaluation integrity and independence are partially guaranteed through the use of some ethical and / or professional standards for evaluators**


While there are some professional and / or ethical standards for evaluators, these may not be comprehensive or universally enforced. Alternatively, these standards do not specifically apply to the evaluation function, creating potential independence, ethical or integrity gaps and potential risks safeguarding evaluation participants' rights, interests, safety or privacy.

**Evaluation integrity and independence are guaranteed through the use of ethical and professional standards for evaluators**


Government evaluators are required to adhere to specific professional standards and/or accreditations, ensuring high criteria for integrity, independence, and competence in evaluation roles. There may be designated job category or accreditation process for government evaluators within the civil service, emphasising the importance of expertise and professionalism in evaluation roles. Professional standards or accreditations outline clear expectations for evaluators regarding integrity, independence, and competence, encompassing knowledge of evaluative methods and culturally appropriate approaches.

**Comments, justification, observations**

**Evaluation autonomy promotion**

 Opportunity for improvement

 Developing

 At standard

**The functional autonomy of internal and external evaluations is not guaranteed**

There are no measures in place to ensure the functional independence of internal or external evaluations. Internal evaluation teams have limited autonomy in resource allocation or decision making. Limited promotion of autonomy undermines the independence and objectivity of evaluations, potentially compromising the credibility and trustworthiness of findings.

**The functional autonomy of internal and external evaluations is partially guaranteed**

The functional autonomy of evaluations is partially promoted by some measures to safeguard either external or internal evaluation processes. This may include commissioning evaluations to external institutions or actors, providing limited autonomy to internal evaluation teams in resource utilisation or decision making, or implementing professional and ethical standards for evaluators. However, the levels of functional autonomy of evaluations differ, mainly between internal and external evaluations. Partial autonomy, coupled with external influences or constraints, may increase the risk of bias or conflicts of interest in evaluation activities, undermining the objectivity and credibility of findings.


**The functional autonomy of internal and external evaluations is systematically guaranteed**

The functional autonomy of evaluations is promoted through multiple measures to safeguard both external and internal evaluation processes. This includes commissioning evaluations to external institutions or actors, providing internal evaluation teams with autonomy in resource utilisation, decision making on which policies to evaluate, control over staff recruitment, and the flexibility to set their own evaluation plans within their mandate. Strong promotion of evaluation autonomy enhances the independence and credibility of evaluations, mitigating potential conflicts of interest and bias.

**Comments, justification, observations**

**Pillar 2** Dimension 4: Quality & Standards

Criteria: Evaluation quality assurance

 Opportunity for improvement


**Some evaluations are conducted and assessed against pre-defined quality criteria**

Evaluations receive minimal or no external scrutiny, and there are no pre-defined quality criteria established to assess evaluation quality systematically. This limited external scrutiny and absence of quality criteria may raise concerns about the credibility and reliability of evaluation findings. Limited peer review and absent quality criteria may erode confidence in evaluation findings, as stakeholders may question the rigor and reliability of evaluations without external validation or systematic quality assessment.

 Developing

**Some evaluations are conducted and assessed against pre-defined quality criteria**

While some evaluations undergo external scrutiny, the process may not be applied universally to all evaluations. Additionally, pre-defined quality criteria for evaluations exist, but they are not published, limiting transparency and accessibility in assessing evaluation quality. Variability in the application of peer review processes and quality criteria may lead to inconsistencies in evaluation quality assessment, raising concerns about credibility and reliability in evaluation practice.

 At standard


**Most evaluations are conducted and assessed against pre-defined quality criteria**

Most evaluations, especially strategic ones, undergo thorough external scrutiny through for example, a peer review system, ensuring rigorous assessment of their methodology, findings, and conclusions. Additionally, predefined quality criteria for evaluations are established and published, providing transparency and accountability in the evaluation process. Comprehensive peer review and published quality criteria enhance the credibility and reliability of evaluation findings, instilling confidence in stakeholders regarding the validity and robustness of the evaluations.

Comments, justification, observations

**Pillar 2** Dimension 5: Institutional Capacities & Skills

Public sector evaluation skill building

 Opportunity for improvement


**Lagging public sector evaluation skills for the delivery of basic evaluation functions**

Measures to acquire and retain employee skills for policy evaluation are absent from government human resource strategies (e.g., no specific job families for evaluation specialists, no recognition or understanding of evaluationspecific competencies). The pool of evaluation specialists in government is limited, and there are limited opportunities for upskilling or training. Non-evaluation specialists have limited understanding of basic principles of evaluation. A limited number of evaluations is performed given the lack of human resources to conduct and commission them, and evaluations which are conducted are of limited quality.

 Developing

**Partially sufficient public sector evaluation skills for the delivery of evaluation functions**

Some measures are in place to acquire and retain employee skills for policy evaluation, but they are not comprehensive. Evaluation specialists are not recognised as a specific job family, and there is a limited understanding of the competencies upon which they can be hired and trained in human resource strategies. Some training and upskilling opportunities are provided, but these mainly target existing evaluation specialists or senior managers. There is limited collaboration with academia and / or the private sector to build skills for conducting and commissioning evaluations. Existing evaluation specialists face some skills gaps, limiting the quality and effectiveness of evaluations.

 At standard


**Sufficient public sector evaluation skills for the delivery of all evaluation functions**

Government human resource strategies include evaluation-specific skills and competency frameworks. Evaluation specialists may be hired through specific contractual modalities and may have their own job families within the civil service, signalling the importance given to this profession in government. Training and upskilling opportunities are available for evaluation specialists and non-specialists alike, and at all levels (senior and junior). Expertise for training is often sourced through collaboration with academia and the private sector. This fosters a culture of continuous improvement in capacities to conduct and commission evaluations.


Comments, justification, observations

**Pillar 2** Dimension 5: Institutional Capacities & Skills

**Data availability assurance**

 Opportunity for improvement

 Developing

 At standard

**Limited or absent availability of data for evaluation**

Government provisions to guarantee and regulate the access and use of data for evaluation are limited, and there is no formal recognition of the role of data in supporting evidence-informed policymaking. Existing provisions are inconsistent, leading to fragmentation in data access principles across ministries and fields of research. There is a lack of clear mechanisms for ensuring privacy protection and confidentiality while facilitating data access for evaluation and research. This may expose sensitive data to privacy risks and breaches, undermining public trust and confidence in data use for evaluation.

**Partial availability of data for policy evaluation**


There are some provisions (e.g. laws or regulations, guidelines) regulating data access and use for research, analytical, or evaluation purposes; but these are not framed in a comprehensive national strategy. Instead, they are defined by each ministry/agency. Not all ministries/agencies have adopted such provisions, and there are some differences in the existing ones, including with regard to privacy and confidentiality protection. Alternatively there are provisions offering broad access to data for evaluation, but these are not accompanied by provisions to safeguard privacy or confidentiality.

**High availability of data for policy evaluation**


There is a national government-wide strategy and related provisions (e.g. laws or regulations, guidelines) in place to promote the availability and accessibility of data for evaluation and evidence-informed policymaking. The strategy is clearly defined and publicly available, outlining objectives, priorities, and action plans for enhancing data availability and accessibility. The strategy includes measures to ensure data is timely, accessible, disaggregated, and reusable for research, analytical, and evaluation purposes. Provisions are made to safeguard privacy and confidentiality while facilitating data access for evaluation and policymaking.

**Comments, justification, observations**

**Institutional resource provision**

 Opportunity for improvement

 Developing

 At standard

**Insufficient institutional resources to fulfil basic evaluation functions**

The evaluation champion(s) - if there is one - lacks sufficient staffing and financing, resulting in significant constraints on its ability to fulfil basic evaluation functions. The availability of resources for evaluation units at the ministry/agency level is also very limited. Financing sources are primarily reliant on budget allocations from the hosting institution, on a shortterm basis. Evidence of severe budget constraints or resource limitations hampers the unit's ability to initiate, conduct, or complete evaluation activities, leading to significant delays or disruptions, and poor quality, undermining the credibility and effectiveness of the evaluation function.

**Sufficient institutional resources to fulfil some evaluation functions**

The evaluation champion(s) faces some staffing and financing shortages to fully ensure all its evaluation functions. There are also important differences and gaps in the levels of staffing and financing in evaluation units at the ministry/agency level. This may result in capacity constraints and workload challenges. Financing sources may include budget allocations from the hosting institution, but dedicated budget allocations earmarked for the unit(s) may be limited or inconsistent. There may be evidence of budget constraints or resource limitations impacting the unit's ability to carry out evaluation activities effectively, such as delays or compromises in project delivery / quality.


**Adequate institutional resources to fulfil all evaluation functions**

The size of the workforce in full-time equivalents (FTEs) of the evaluation champion(s) and ministry/agency evaluation unit(s) meets or exceeds the required capacity to carry out all evaluation functions. Financing sources include dedicated budget allocations earmarked for the unit(s) or budget allocations from the hosting institution that adequately support evaluation activities. Budget allocations ensure stable funding for evaluation activities, reducing reliance on ad-hoc or unpredictable funding sources. There is evidence that the evaluation unit(s) has the necessary financial resources to cover operational costs, including staff salaries, training, and other expenses related to evaluation activities.


**Comments, justification, observations**

**Pillar 3 Dimension 6: Use in decision making**

**High-level guidance establishment**

 Opportunity for improvement

 Developing

 At standard

**No high-level guidance exists for framing policy evaluation across the government**

There are no documented frameworks, agendas, plans, or guidelines specifically outlining the process, methodologies, or requirements for conducting policy evaluation across government entities. This creates a risk of inconsistency and variability in evaluation processes and methodologies across government institutions. In addition, the absence of legal or policy requirements for government institutions to conduct or use evaluations may result in ad-hoc evaluation practices, leading to inefficient use of resources, and missed opportunities for learning and improvement.

**Some high-level guidance for framing evaluation is present**


There are documented frameworks, agendas, plans, or guidelines providing general direction or recommendations for conducting policy evaluation in some policy areas or line ministries. In addition, there may be some legal or policy requirements for government institutions to conduct and use evaluations, although they may be limited in scope or detail, reflecting partial legal recognition or support for evaluation practices.

**Comprehensive high-level guidance framing policy evaluation**


There are well-defined frameworks, agendas, plans, or guidelines offering detailed and structured guidance on the process, methodologies, and standards for conducting policy evaluation across government entities. These are designed to allow flexibility in adapting evaluation processes to unique contexts or evolving needs. These may supported by comprehensive legal or policy requirements for government institutions to conduct and use evaluations.

**Comments, justification, observations**

**Evaluation results incorporation**

 Opportunity for improvement

 Developing

 At standard

**Limited incorporation of evaluation results in decision making**

Evaluation findings are rarely considered in the development of government programmes, strategic plans, in programme management, or in budgetary decisions. Instances of using evaluation findings tend to be ad-hoc rather than part of a structured process, indicating a lack of systematic incorporation of evaluation results into decision-making and management frameworks. No practices have been adopted to ensure policy evaluations support budgetary decision making, such as performance budgeting, spending reviews or ex ante appraisals of capital and/or current expenditure.

**Moderate incorporation of evaluation results in decision making**

Evaluation findings are occasionally considered in the development of government programmes and strategic plans, in programme management, or budgetary decisions. The mechanisms for this purpose lack consistency and are not applied in all key policymaking and budgetary decision process. There is a continued risk of suboptimal decision making due to inconsistent incorporation of evaluation findings. Some practices such as performance budgeting or spending reviews have been adopted and draw from evaluation findings to support budgetary decision making; indicating a limited yet growing awareness of the value of evaluation and its potential to improve resource allocation.

**Extensive incorporation of evaluation results in decision making**

Evaluation findings consistently inform the development of government programmes and strategic plans, programme management, and budgetary decisions; thanks to the use of mechanisms to integrate evaluation results into decision making (e.g. performance budgeting, spending reviews or ex ante appraisals of capital and/or current expenditure). This has enhanced accountability and transparency in decision-making processes, leading to improved public trust and confidence; without creating rigidity in decision-making processes.

**Comments, justification, observations**

**Pillar 3 Dimension 6: Use in decision making**

**Follow-up mechanism**

● Opportunity for improvement

● Developing

● At standard

**Limited follow-up of evaluation results and recommendations**

There is no established mechanism for addressing evaluation findings or implementing recommendations. While evaluation results are sometimes discussed or acknowledged by programme owners or managers, there is no systematic approach to follow-up on their implementation. This leads to a lack of accountability or oversight in ensuring that evaluation findings lead to action or improvements, and generate a risk that evaluation findings are overlooked, leading to stagnation or decline in programme effectiveness.

**Partial follow-up of evaluation results and recommendations**

Evaluation follow-up mechanisms exist (e.g. management response mechanisms) but they are not consistently applied across all evaluations. Consequently, some evaluation findings lead to action, but others may be overlooked or addressed inadequately. There may be delays in responding to evaluation recommendations, or follow-up actions may lack transparency. Strategic evaluations are sometimes discussed at the highest level (e.g., Council of Ministers), but this is done on a very selective basis. There are still missed opportunities to address critical evaluation findings or implement necessary changes, undermining stakeholders' appreciation of the importance of evaluation.

**Systematic follow-up of evaluation results and recommendations**

Clear procedures for responding to evaluation findings (e.g. management response mechanisms, follow-up committees) are in place, which ensure that a response is given to evaluation findings by management at the level of line ministries, and a significant number of strategic evaluations are discussed at the highest level such as at the Council of Ministers. This allows to take timely actions to address recommendations or issues raised in evaluations, and regularly monitor and track the follow-up actions to ensure accountability and ensure improvement in programme delivery and outcomes.

**Comments, justification, observations**

**Pillar 3 Dimension 7: Accessibility of findings and results**

**Evaluation findings availability**

● Opportunity for improvement

● Developing

● At standard

**Limited accessibility of evaluation results**

Evaluation reports are mostly only made available to internal government stakeholders or primary users. There is no broader dissemination and access is not possible for the public or any other stakeholder. There is no centralised repository or database of evaluations. As a result, the opportunities for learning based on evaluation findings are limited and the government's transparency and accountability reputation may be tainted.

**Partially accessible evaluation results**

Evaluations are either automatically published in some cases or are accessible to the public upon request through established channels. There may be a designated process for requesting evaluation reports, as well as a centralised repository or database with a limited number of publicly available evaluations. While this ensures some level of transparency and accountability, it also increases the barriers to access evaluation findings. While this approach allows for flexibility in managing evaluation report dissemination while still meeting some transparency standards, the use in decision making is reduced given the barriers and delays in accessing reports.


**Fully accessible evaluation results**

Evaluation results are routinely published and accessible to the public without the need for specific requests. This may be facilitated by the existence of a designated platform or database where evaluation reports are posted for easy access. Transparency and openness are prioritised in disseminating evaluation findings to a wide audience of policymakers, researchers and the general public. Open dissemination facilitates knowledge sharing and learning from evaluation outcomes, leading to potential improvements in programme design and implementation. Some level of curation is used to limit privacy concerns and ensure searchability of documents.


**Comments, justification, observations**

**Pillar 3 Dimension 7: Accessibility of findings and results**

**Evaluation evidence presentation**

 Opportunity for improvement

 Developing

 At standard

**Limited tailoring and accessibility of evaluation results for users**

Tailored formats for publication, as well as tailored communication strategies are seldom used and deployed. Instead, evaluations either remain unpublished or are published in their original format which is not fit for purpose for broader stakeholder groups beyond immediate users. There is no strategic guidance established for the dissemination of evaluation findings and the format of content by different user groups.

**Partially tailored and accessible evaluation results for users**


A limited number of evaluations is published with some level of tailoring, making the findings more accessible by users. Tailored communication strategies are used in some cases, but are not fully widespread due to resource constraints or other factors. While tailored formats for some evaluations still enhance usability and accessibility, there are still some missed opportunities for broader numbers of stakeholders to effectively capitalise on evaluation findings and evidence.

**Regularly tailored and accessible evaluation results for intended users**


Evaluation findings are synthesised and presented in formats such as policy briefs, executive summaries, or infographics. Tailored communication strategies are adopted to disseminate evaluation results to well identified users or stakeholders. Accessible formats enhance understanding and usability of evaluation findings, facilitating their uptake and application. These also ensure that evaluation results are readily available and comprehensible to diverse audiences, including policymakers, practitioners, and the public.

**Comments, justification, observations**

**Evidence synthesis use**

 Opportunity for improvement

 Developing

 At standard

**Limited or no aggregation of evaluation findings for systematic assessment**

Evidence synthesis methodologies are rarely or never used by evaluation champions to aggregate evaluation findings. Instead they rely primarily on individual evaluation reports without systematically aggregating or assessing them. Without synthesis, decision makers miss valuable insights and trends that could inform more effective policies and programmes, and the reliance on individual evaluations increases the risk of biased or incomplete decision making as broader patterns or contradictory evidence may be overlooked.

**Occasional aggregation of evaluation findings for systematic assessment**

Some evidence synthesis methodologies are used sporadically or inconsistently by evaluation champions to aggregate evaluation findings. Synthesis efforts may be limited to specific evaluations or contexts, rather than encompassing a broader range of evidence sources. The use of evidence synthesis is not systematically integrated into decision-making processes or championed as a standard practice. This may result in missed opportunities to uncover broader patterns or trends across evaluations, limiting the depth of insight gained.

**Regular aggregation of evaluation findings for systematic assessment**

Evaluation champions consistently use a range of evidence synthesis methodologies to aggregate evaluation findings and assess them systematically. Synthesis efforts consider a wide range of evaluations to provide a comprehensive understanding of the effectiveness of policies or programmes, which strengthens the evidence base of decision making.

**Comments, justification, observations**

# The OECD Recommendation on Public Policy Evaluation

## Assessment Grid



Pillars	Dimensions	Criteria	Opportunity for improvement	Developing	At standard	
 <b>Institutionalisation</b>	Execution & Agency	Evaluation Champion Designation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Institutional Responsibility Definition	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
	Learning & Accountability culture	Knowledge Broker Promotion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Legislative Review Opportunity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
 <b>Quality</b>	Design, Planning & Management	Early Evaluation Planning & Provision	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Evaluation Adaptation Design	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Stakeholder Engagement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
	Quality & Standards	Evaluation Guidelines Development	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Professional Standards Adoption	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Evaluation Autonomy Promotion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Evaluation Quality Assurance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
	Institutional Capacities & Skills	Public Sector Evaluation Skill Building	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Data Availability Assurance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Institutional Resource Provision	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
	 <b>Impact</b>	Use in Decision-making	High-Level Guidance Establishment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
			Evaluation Results Incorporation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Follow-up Mechanism Establishment			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Accessibility of Findings & Results		Evaluation Findings Presentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Evaluation Evidence Tailoring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
		Evidence Synthesis Utilisation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

# Annex C Evaluation guidelines repository

## Australia

- Commonwealth Evaluation Policy and Toolkit (<https://www.finance.gov.au/government/managing-commonwealth-resources/planningand-reporting/commonwealth-performance-framework/evaluation-commonwealth-rmg130>)
- Australian Government Guide to Policy Impact Analysis (<https://oia.pmc.gov.au/resources/guidance-impact-analysis/australian-governmentguide-policy-impact-analysis>)

## Austria

- WFA Handbuch (<https://oeffentlicherdienst.gv.at/wp-content/uploads/2023/01/HandbuchWirkungsorientierte-Folgenabschaetzung.pdf>)
- The Austrian Platform for Research and Technology Policy Evaluation (fteval) Standards (<http://repository.fteval.at/id/eprint/387>)

## Belgium

## Canada

- Policy on Results (<https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=31300>)
- Directive on Results (<https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=31306>)
- Guide to Rapid Impact Evaluation (<https://www.canada.ca/en/treasury-boardsecretariat/services/audit-evaluation/evaluation-government-canada/guide-rapid-impacetevaluation.html>) •
- Integrating Gender-Based Analysis Plus into Evaluation: A Primer (<https://www.canada.ca/en/treasury-board-secretariat/services/auditevaluation/evaluation-government-canada/gba-primer.html>)
- Policy on Results: What is Evaluation? (<https://www.canada.ca/en/treasury-boardsecretariat/services/audit-evaluation/evaluation-government-canada/policy-results-whatevaluation.html>)

## Chile

- Evaluación Ex-Post Conceptos y Metodologías ([https://www.dipres.gob.cl/598/articles266586\\_doc\\_pdf.pdf](https://www.dipres.gob.cl/598/articles266586_doc_pdf.pdf))
- Evaluación Focalizada de Ámbito EFA ([https://www.dipres.gob.cl/598/articles266587\\_doc\\_pdf.pdf](https://www.dipres.gob.cl/598/articles266587_doc_pdf.pdf))
- Proceso de evaluación ex ante Diseño de Programa ([https://www.dipres.gob.cl/598/articles-285500\\_C\\_Instructivo\\_ExAnte\\_2023.pdf](https://www.dipres.gob.cl/598/articles-285500_C_Instructivo_ExAnte_2023.pdf))
- Regulation that sets criteria and procedures for the evaluations of social programmes (<https://bcn.cl/2z3xc>)

## Colombia

- Guía Metodológica (<guc3ada-para-la-evaluac3b3n-de-pp.pdf> (wordpress.com))
- Guía Metodológica para el Seguimiento y Evaluación de Políticas Públicas (<https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Cartilla%20Guia%20para%20Seguimiento%20y%20Evaluaci%C3%B3n%20Ago%2013.pdf>)

## **Costa Rica**

- Manual of Evaluation for public interventions (<https://www.mideplan.go.cr/herramientasmetodologicas-en-evaluacion>)
- Guideline for theory programmes (<https://www.mideplan.go.cr/herramientasmetodologicas-en-evaluacion>)
- Guideline for evaluation products (<https://www.mideplan.go.cr/herramientasmetodologicas-en-evaluacion>) • Guideline for evaluation use (<https://www.mideplan.go.cr/herramientasmetodologicas-en-evaluacion>)
- Guideline for participatory evaluations (<https://www.mideplan.go.cr/herramientasmetodologicas-en-evaluacion>)

## **Czechia**

- Metodický Pokyn Pro Oblast Indikátorů, Evaluací A Publicity V Programovém Období 2021-2027(<https://www.dotaceeu.cz/getattachment/61c005d1-d800-4e2a-b9a8-b7eb82ca7c3b/MP-Indikatory-Evaluace-Publicita-2021-2027.pdf.aspx?lang=csCZ&ext=.pdf>) •
- Evaluator's Guide ([https://www.dotaceeu.cz/getmedia/dd31644f-b766-4cd4-8a80-d2b0296238b4/EVALUATOR%E2%80%99S-GUIDE\\_2020.pdf.aspx](https://www.dotaceeu.cz/getmedia/dd31644f-b766-4cd4-8a80-d2b0296238b4/EVALUATOR%E2%80%99S-GUIDE_2020.pdf.aspx))

## **Denmark**

### **European Commission**

- Better regulation: guidelines and toolbox ([https://commission.europa.eu/law/law-makingprocess/planning-and-proposing-law/better-regulation/better-regulation-guidelines-andtoolbox\\_en](https://commission.europa.eu/law/law-makingprocess/planning-and-proposing-law/better-regulation/better-regulation-guidelines-andtoolbox_en))

## **Estonia**

- A rule of good lawmaking and normative regulation (Hea õigusloome ja normitehnika eeskiri) (<https://www.riigiteataja.ee/akt/129122011228>)
- Methodology of Impact Assessment (Mõjude hindamise meetodika) (<https://www.riigikantselei.ee/valitsuse-too-planeerimine-ja-korraldamine/mojudehindamine>)

## **France**

- Faire de l'évaluation des politiques publiques un véritable outil de débat démocratique et de décision (<https://www.conseil-etat.fr/Media/actualites/documents/2020/08-aout/etude2020-evaluation-des-politiques-publiques.pdf>)

## **Greece**

## **Hungary**

## **Iceland**

## **Ireland**

- Public Spending Code (<https://www.gov.ie/en/publication/public-spending-code/>)

## Italy

- Criteria and methodologies for the definition of processes and activities for the analysis and evaluation of expenditure (<https://area.rgs.mef.gov.it/canali/498/bacheca>)
- Guidelines on ex post and ongoing evaluations (Linee guida per attività valutative ex post e in itinere)  
(<https://www.bing.com/ck/a?!&p=759bd808e1e50c55JmltdHM9MTY4Njc4NzlwMCZpZ3VpZD0yYTdiOGUzNC1iYzViLTZkM2QtMDQyNS05ZDA1YmQ1YzZjM2MmaW5zaWQ9NTI5NQ&ptn=3&hsh=3&fclid=2a7b8e34-bc5b-6d3d-0425-9d05bd5c6c3c&psq=NUVAP+2015+linee+guida+piani+di+valutazione&u=a1aHR0cHM6Ly93d3cudmFsdXRhemlVjbmVjb2VzaW9uZS5pdC9kb2MvaW5jb250cmkvMjAxN190b3RhLU5VVkFQX0xpbmVILUd1aWRhLXZhbHV0YXppb25pLWV4LXBvc3RfaW4taXRpbmV5ZS5wZGY&ntb=1>)
- Guide on Impact Analysis and Verification (Guida all'analisi e alla verifica dell'impatto della regolamentazione)  
([https://www.gazzettaufficiale.it/atto/serie\\_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2018-04-10&atto.codiceRedazionale=18A02428&elenco30giorni=false](https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2018-04-10&atto.codiceRedazionale=18A02428&elenco30giorni=false))
- Guidelines for evaluation of public investments (Linee guida per la valutazione degli investimenti in opere pubbliche)  
(<https://www.mit.gov.it/ricerca/Linee%20guida%20valutazione%20opere%20pubbliche>)

## Japan

- Policy Evaluation Implementation Guidelines  
([https://www.soumu.go.jp/main\\_content/000354069.pdf](https://www.soumu.go.jp/main_content/000354069.pdf))
- Guidelines for Publication of Information on Policy Evaluation  
([https://www.soumu.go.jp/main\\_content/000556224.pdf](https://www.soumu.go.jp/main_content/000556224.pdf))
- Implementation Guidelines for Policy Evaluation of regulations  
([https://www.soumu.go.jp/main\\_content/000499513.pdf](https://www.soumu.go.jp/main_content/000499513.pdf))
- Target Management-based Policy Evaluation Implementation Guidelines  
([https://www.soumu.go.jp/main\\_content/000556226.pdf](https://www.soumu.go.jp/main_content/000556226.pdf))
- Implementation Guidelines for Policy Evaluation Pertaining to Special Taxation Measures  
([https://www.soumu.go.jp/main\\_content/000556225.pdf](https://www.soumu.go.jp/main_content/000556225.pdf))

## Latvia

- Policy Planning Handbook  
([https://www.pkc.gov.lv/sites/default/files/inlinefiles/pkc\\_rokasgramata\\_090316\\_web.pdf](https://www.pkc.gov.lv/sites/default/files/inlinefiles/pkc_rokasgramata_090316_web.pdf))
- Practical Recommendations of the European Union Investment Evaluation  
(<https://m.esfondi.lv/upload/praktiskie-izvertesanas-ieteikumi-.pdf>)

## Lithuania

- Methodology on Strategic Governance  
(<https://eseimas.lrs.lt/portal/legalAct/lt/TAD/5e3aa191a8e511eb98ccba226c8a14d7/asr>)
- Methodological Guidance on Assessment of EU Investment Funds  
(<https://2014.esinvesticijos.lt/lt/dokumentai//es-fondu-investiciju-vertinimo-metodinesgaires-1>)

- Evaluation Standards for the EU structural assistance to Lithuania ([https://2014.esinvesticijos.lt/media/force\\_download?url=/uploads/main/documents/docs/250\\_74b4a05aa8d378194e7dc28bff9930f8.pdf](https://2014.esinvesticijos.lt/media/force_download?url=/uploads/main/documents/docs/250_74b4a05aa8d378194e7dc28bff9930f8.pdf))
- Methodology on impact assessment of envisaged legal regulation (<https://eseimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.205970/asr>)
- Methodology on ex post evaluation of the existing legal regulation (<https://eseimas.lrs.lt/portal/legalAct/lt/TAD/6cb6bfd4b15311eba6328c92adabc234?positionInSearchResults=14&searchModelUUID=b1579e87-5b04-48bf-ab39-11afef4b3b3c>)

## Luxembourg

## Mexico

- General Guidelines for the Evaluation of Federal Programs of the Federal Public Administration ([https://www.coneval.org.mx/rw/resource/coneval/eval\\_mon/361.pdf](https://www.coneval.org.mx/rw/resource/coneval/eval_mon/361.pdf))
- Terms of Reference for types of evaluation ([https://www.coneval.org.mx/Evaluacion/MDE/Paginas/Evaluaciones\\_Disenio.aspx](https://www.coneval.org.mx/Evaluacion/MDE/Paginas/Evaluaciones_Disenio.aspx))
- Annual Evaluation Program ([https://www.coneval.org.mx/Evaluacion/NME/Documents/PAE\\_2023.pdf](https://www.coneval.org.mx/Evaluacion/NME/Documents/PAE_2023.pdf))
- Terms of Reference for the Performance Evaluation of the Federal Contribution Funds in the Federal Entities ([https://www.coneval.org.mx/Evaluacion/ERG33/paginas/metodologias\\_evaluacion.aspx](https://www.coneval.org.mx/Evaluacion/ERG33/paginas/metodologias_evaluacion.aspx))
- Aspects to consider for the preparation of budget programme diagnoses ([https://www.coneval.org.mx/Evaluacion/MDE/Paginas/Diagnostico\\_Programas\\_Nuevos.aspx](https://www.coneval.org.mx/Evaluacion/MDE/Paginas/Diagnostico_Programas_Nuevos.aspx))

## Netherlands

- Regeling Periodiek Evaluatieonderzoek 2022 (<https://zoek.officielebekendmakingen.nl/stcrt-2022-19587.html>)
- Toolbox Beleidsevaluaties (<https://evaluaties.rijksfinancien.nl/toolbox>) New Zealand • Evidence and Evaluation (<https://www.dpmc.govt.nz/our-programmes/policyproject/navigator/evidence-and-evaluation>)

## New Zealand

- Evidence and Evaluation (<https://www.dpmc.govt.nz/our-programmes/policyproject/navigator/evidence-and-evaluation>)

## Norway

- Guide for CBA (<https://dfo.no/fagomrader/utredning-og-analyse-av-statlige-tiltak/veilederisamfunnsokonomiske-analyser>)
- Guidance notes on the Instructions for Official Studies of Central Government Measures "Instructions for Official Studies" ([https://dfo.no/sites/default/files/fagomr%C3%A5der/Utreddingsinstruksen/Guidance\\_Notes\\_on\\_the\\_Instructions\\_for\\_Official\\_Studies.pdf](https://dfo.no/sites/default/files/fagomr%C3%A5der/Utreddingsinstruksen/Guidance_Notes_on_the_Instructions_for_Official_Studies.pdf))
- Guidelines to carry out evaluations (<https://dfo.no/fagomrader/etatsogvirksomhetsstyring/evaluering/veiledere-til-evaluering>)
- Strategic and systematic use of evaluation in public management (<https://dfo.no/fagomrader/etats-og-virksomhetsstyring/evaluering/veiledere-til-evaluering>)
- Evaluation of laws (<https://dfo.no/fagomrader/etatsogvirksomhetsstyring/evaluering/veiledere-til-evaluering>)

## Poland

- Guidance for evaluation cohesion policy 2014-2020 (<https://www.funduszeuropejskie.gov.pl/strony/o-funduszach/dokumenty/wytyczne-wzakresie-ewaluacji-polityki-spojnosci-na-lata-2014-2020>)
- Guidance for evaluation of cohesion policy 2021-2027 ([https://www.funduszeuropejskie.gov.pl/media/110684/wytyczne\\_ewaluacja\\_240822.pdf](https://www.funduszeuropejskie.gov.pl/media/110684/wytyczne_ewaluacja_240822.pdf))

## Portugal

- Portugal 2020 Global Evaluation Plan ([https://www.adcoesao.pt/wpcontent/uploads/01\\_PGA-PT2020\\_Revisao2021\\_dez2021-0002.pdf](https://www.adcoesao.pt/wpcontent/uploads/01_PGA-PT2020_Revisao2021_dez2021-0002.pdf))
- Guidelines for the planning and preparation of Portugal 2020 evaluations ([https://www.adcoesao.pt/wpcontent/uploads/pga\\_orientacoes\\_planeamento\\_preparacao.pdf](https://www.adcoesao.pt/wpcontent/uploads/pga_orientacoes_planeamento_preparacao.pdf))
- Guidelines for monitoring the implementation of the recommendations of the Portugal 2020 evaluations (Follow-up) ([https://www.adcoesao.pt/wpcontent/uploads/pga\\_orientacoes\\_acompanhamento\\_implementation\\_recomendacoes.pdf](https://www.adcoesao.pt/wpcontent/uploads/pga_orientacoes_acompanhamento_implementation_recomendacoes.pdf))
- Portugal 2030 Global Evaluation Plan ([https://www.adcoesao.pt/wp-content/uploads/pga\\_pt2030.pdf](https://www.adcoesao.pt/wp-content/uploads/pga_pt2030.pdf))

## Spain

- Guidelines for the Evaluability Assessment of Public Policies ([https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/01\\_GUIDELINES\\_EVALUABILITY\\_ASSESSMENT.pdf#page=1](https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/01_GUIDELINES_EVALUABILITY_ASSESSMENT.pdf#page=1))
- Guidelines for the design evaluation of Public Policies ([https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/02\\_GUIDELINES\\_DESIGN\\_EVALUATION\\_PRINT.pdf#page=1](https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/02_GUIDELINES_DESIGN_EVALUATION_PRINT.pdf#page=1))
- Guidelines for the implementation evaluation of Public Policies ([https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/03\\_GUIDELINES\\_IMPLEMENTACION\\_EVALUACION\\_PRINT.pdf#page=1](https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/03_GUIDELINES_IMPLEMENTACION_EVALUACION_PRINT.pdf#page=1))
- Guidelines for the outcome and impact evaluation of Public Policies ([https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/04\\_GUIDELINES\\_OUTCOME\\_IMPACT\\_EVALUACION\\_PRINT.pdf#page=1](https://funcionpublica.hacienda.gob.es/dam/es/portalsefp/evaluacion-politicaspUBLICAS/Documentos/Metodologias/04_GUIDELINES_OUTCOME_IMPACT_EVALUACION_PRINT.pdf#page=1))

## Sweden

### Switzerland

- Standards d'évaluation de la Société suisse d'évaluation ([https://www.seval.ch/app/uploads/2020/12/Standards-SEVAL-2016\\_f.pdf](https://www.seval.ch/app/uploads/2020/12/Standards-SEVAL-2016_f.pdf))
- Explications relatives aux standards d'évaluation de la Société suisse d'évaluation ([https://www.seval.ch/app/uploads/2020/12/Standards-SEVAL\\_2016-Explications.pdf](https://www.seval.ch/app/uploads/2020/12/Standards-SEVAL_2016-Explications.pdf))
- Guide de législation (<https://www.bj.admin.ch/dam/bj/fr/data/staat/legistik/hauptinstrumente/gleitff.pdf.download.pdf/gleitf-f.pdf>)

## **United Kingdom**

- Magenta Book (<https://www.gov.uk/government/publications/the-magenta-book>)
- Green Book (<https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2020>)
- Acqua Book  
([https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/416478/aqua\\_book\\_final\\_web.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/416478/aqua_book_final_web.pdf))

## **United States**

- OMB M-20-12: Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices (<https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf>)
- OMB M-18-04: Monitoring and Evaluation Guidelines for Federal Departments and Agencies that Administer United States Foreign Assistance (<https://www.whitehouse.gov/wp-content/uploads/2017/11/M-18-04-Final.pdf>)
- OMB M-21-27: Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans (<https://www.whitehouse.gov/wp-content/uploads/2021/06/M-21-27.pdf>)
- OMB M-19-23: Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance (<https://www.whitehouse.gov/wp-content/uploads/2019/07/m-19-23.pdf>)

# Annex D National evaluation champions

## **Australia**

- Australian Centre for Evaluation

## **Austria**

- Federal Ministry for Arts, Culture, Civil Service and Sports

## **Belgium**

- Ministry of Public Administration

## **Canada**

- Treasury Board Secretariat of Canada (TBS)

## **Chile**

- Ministry of Finance Ministry of Planning Colombia National Planning Department

## **Costa Rica**

- Ministry of Planning and Economic Policy Czechia Governmental Analytical Unit (VAU)

## **Denmark**

## **European Commission**

- Secretariat-General, Dir. A - Strategy, Better Regulation & Corporate Governance, Unit A2 – Evaluation & Impact Assessment

## **Estonia**

- Government Office Ministry of Finance Ministry of Justice

## **France**

## **Greece**

- Centre of Government

## **Hungary**

- Cabinet Office of the Prime Minister, Prime Minister's Office
- Ministry of Finance
- Government Control Office

## **Iceland**

- Office of Prime Minister
- Ministry of Finance
- Ministry of Interior

## **Ireland**

- Department of Public Expenditure

**Italy**

- Presidency of the Council Ministry of Finance

**Japan**

- Administrative Evaluation Bureau - Ministry of Internal Affairs and Communications, (AEB)

**Latvia**

- Department of Cross-Sectoral Coordination - State Chancellery

**Lithuania**

- Office of the Government
- Ministry of Finance
- Government Strategic Analysis Centre (STRATA)
- Ministry of Justice

**Luxembourg****Mexico**

- Ministry of Finance
- National Council for Evaluation of Social Development Policy (Coneval)

**Netherlands**

- Ministry of Finance

**New Zealand**

- Department of the Prime Minister and Cabinet
- Treasury
- Public Service Commission
- Social Wellbeing agency

**Norway**

- Ministry of Finance

**Poland**

- Ministry of Planning

**Portugal**

- PlanAPP
- Ministry of Finance (GPEARl)
- Agency of Development and Cohesion (AD&C)

**Spain**

- State Agency for the Evaluation of Public Policies

**Sweden**

**Switzerland**

- Federal Office of Justice

**United Kingdom**

- HM Treasury Evaluation Task Force

**United States**

- Office of Management and Budget, Office of Performance and Personnel Management

**Source: OECD (2023), Survey on Public Policy Evaluation.**

**Note: Answer from Australia was updated following the establishment of the Australian Centre for Evaluation after the survey was conducted**



For more information, visit our website

<https://www.oecd.org/en/topics/public-policy-monitoring-and-evaluation.html>