

How Can Governments Leverage Policy Evaluation to Improve Evidence Informed Policy Making ?

*Highlights
from an OECD comparative study
2020*



Table of contents

Executive Summary	3
Introduction	5
Definitions, methodology and integration into the policy making chain	10
Why is institutionalisation so difficult to achieve?	13
Fostering an evaluation culture through the promotion of quality and use of evaluation	20
Figures	
Figure 1. Government's main objectives for conducting evaluations	6
Figure 2. Challenges for promoting policy evaluation across government	7
Figure 3. Government's formal definition of policy evaluation	10
Figure 4. Conceptual clusters included in the definition of policy evaluation	11
Figure 5. Legal frameworks guiding policy evaluation across government	14
Figure 6. Availability and content of Policy Frameworks	15
Figure 7. Institutions within the Executive that have competences related to policy evaluation across government	16
Figure 8. Mandate of the Centre of Government for Policy Evaluation	17
Figure 9. Mandate of the Ministry of Finance/Economy/Treasury or equivalent for policy evaluation	18
Figure 10. How do countries ensure the quality of evaluations across government?	21
Figure 11. Building competencies for policy evaluation	22
Figure 12. How do countries promote the use of evaluation?	24
Figure 13. Making the results of evaluation public	25
Figure 14. How are policy evaluation findings used into the budget cycle?	26
Boxes	
Box 1. Components of the policy evaluation system	8
Box 2. The Policy chain cycle	12

Executive Summary

Policy evaluation contributes fundamentally to sound public governance. It can help governments improve the design and implementation of public policies that, in turn, can lead to prosperity for their country and well-being for citizens. Policy evaluation contributes to promoting public accountability, learning and increased public sector effectiveness through improved decision-making.

The report provides a broad analysis across 42 countries of the institutionalisation of policy evaluation, its quality and use. This systemic perspective allows for a full discussion of how evaluation can contribute to the policy cycle, as well as an overview of the policy tools that rely upon evaluation, such as budgeting and regulation. A sound institutional set-up can help align isolated and unplanned evaluation efforts into more formal and systematic approaches. Promoting both better quality and use can have a greater impact in ensuring that evaluation fully achieves its purpose.

Generally, countries show strong commitment to policy evaluation. Some countries have embedded policy evaluations in their constitutions, and around two-thirds of responding countries have developed some kind of legal framework for policy evaluation. Similarly, most countries have adopted guidelines on policy evaluation applicable across government.

The term “evaluation” can cover a range of practices. Around two-thirds of the countries surveyed have at least one definition of evaluation. While these definitions of policy evaluation often reflect the specificities of a country’s institutional setting, common concepts can be found across definitions, in particular regarding *what should be measured* (policies, programmes, plans, reforms), *why an evaluation should be conducted* (aims), and *when* (ex-ante or ex-post).

Most countries face significant challenges in promoting policy evaluation across government. These challenges are mainly related to the limited use of evaluation results in policy-making, the absence of a coherent whole-of-government strategy for policy evaluation, and the lack of human resources – whether in terms of skills, capacity or capability.

Sound institutional set-ups can provide incentives to ensure that evaluations are effectively conducted. They can promote transparency and accountability in the management of evaluations, and contribute to improving both the comparability and consistency of results, as well as the use of results in policy processes. For this reason, countries are using a range of institutional approaches to anchor policy evaluations in their governance apparatus. They may also co-ordinate their approaches and frameworks for policy evaluation with those related to evidence and data governance.

In addition to legal frameworks, about half of the countries have also developed policy frameworks that provide principles and strategic direction for evaluations. Such policy frameworks provide high-level guidance and clarity for institutions by outlining overarching best practices and goals.

The centre of government provides strategic direction for policy evaluation in two-thirds of the countries surveyed. As such, the centre plays a crucial role in embedding a whole-of-government approach to policy evaluation and it often has the broadest mandate to do so. Ministries of finance also have significant responsibility in about six countries out of ten. These results show the strong economic impetus for policy

evaluation and the close connection between incentives to enhance the quality of public expenditure and incentives to deliver results. Ministries of public sector reform and planning also play a role in some countries, along with autonomous agencies. Agencies are often a source of good evaluation practices, particularly in the Nordic countries, even if their role remains decentralised. They are generally well placed to conduct independent, transparent and accountable policy evaluations.

Co-ordination mechanisms such as commissions and integrated cross-departmental services, when they exist, can help strengthen evaluation systems, for example in terms of staffing and capabilities. Locating the office in charge of policy evaluation close to political decision-making power may allow it to be more effective in commissioning policy evaluations and following up on commitments by ministries. At the same time, ensuring that evaluation systems are independent, transparent and accountable can help bolster citizens' trust in the results.

Policy evaluation can only be truly effective if it is of high quality and its results are used. Quality control and quality assurance are key to ensuring the robustness of evaluations. Standards play an important role in quality assurance, and guidelines exist in three-quarters of the countries surveyed. Other quality control mechanisms, such as peer review, systematic reviews, and competency requirements for evaluators are relatively common. Up to half of the countries organise training for evaluators, and a majority recognise the importance of developing evaluator competences.

While using the results of evaluations is a challenge, it is paramount to achieving impact. Countries are relying on a range of organisational and institutional mechanisms to promote their use and to create a marketplace for evaluations. Some of these measures consist of a co-ordination platform (in about one-third of the sample) or a management response (in one-quarter of the sample). Rating and grading systems are also used to a limited extent. Finally, most mechanisms to promote skills and competences are aimed at evaluators and managers, rather than improving the capacity of policy makers and decision makers to use evidence.

The heterogeneity of country approaches suggests that the creation of an evaluation marketplace depends on the local political and cultural context. Evaluation can also be embedded into policy planning and policy-making processes. Half of the countries incorporate evaluation results into the budget cycle. In particular, many OECD countries use spending reviews. The area of regulatory policy is also one where the use of evaluation is well developed, with significant requirements for evaluation embedded in the regulatory impact assessment process.

The role of institutions outside of the executive branch remains limited, both in the promotion of quality and use of evaluations, beyond their involvement in the budgetary cycle – although Supreme Audit Institutions are a key actor in terms of the supply of evaluations.

Introduction

Policy evaluation matters

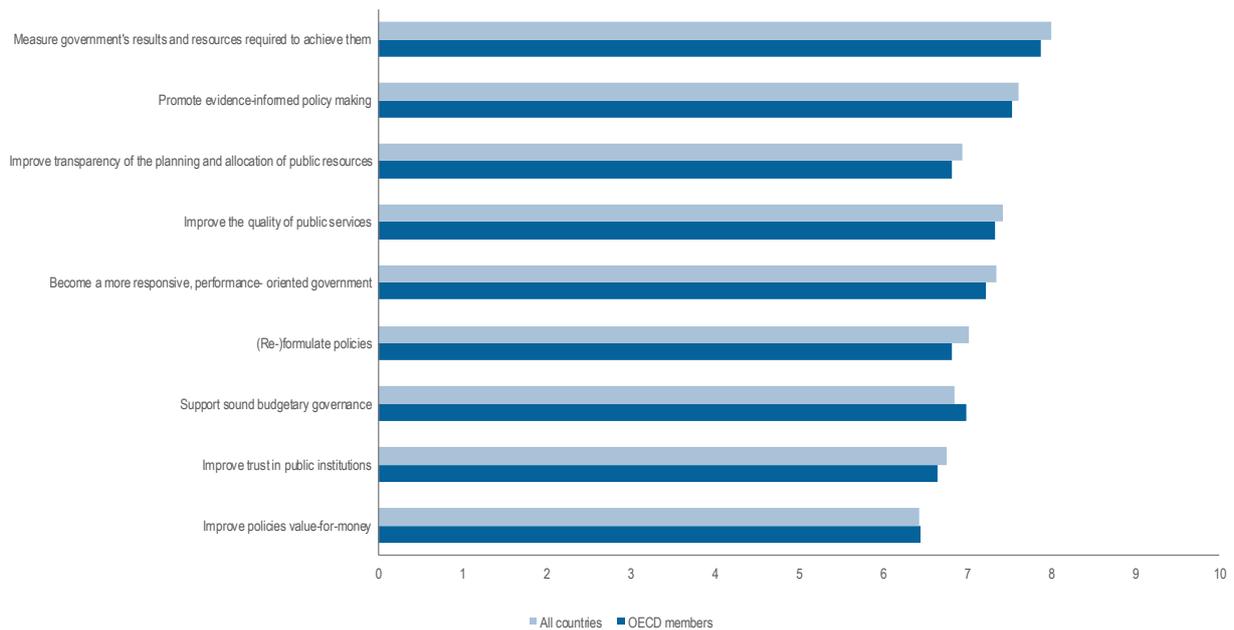
In modern societies, governments do not operate in a vacuum, as they have to face and justify their decisions to a broad array of stakeholders, including citizens and businesses at the domestic level, international investors, or the international community in some cases. Policy evaluation is critical to ensure that policies are actually improving outcomes by bringing an understanding of what works, why, for whom, and under what circumstances. It connects policies, outcomes, policymakers' decisions and main actors: governments and citizens to achieve high levels of well-being, and ensure progress towards government goals, or wide international commitments, such as the Sustainable Development Goals.

Policy evaluation is a core instrument of sound public governance that can contribute to avoiding policy capture, fostering learning and promoting accountability. Governments are asked to demonstrate that their decisions and policies are informed by evidence, that they set realistic expectations about various policy choices, and that they spend public resources adequately. In this sense, policy evaluation facilitates **learning** as it helps us to understand why and how a policy was or has the potential to be successful or not by informing policy makers about the reasons and causal mechanisms leading to policy success or failure. It contributes to the quality of decision-making by providing insights on how to improve links between policy formulation, implementation and outcomes. At the same time, policy evaluation has the potential to improve value for money, **accountability** and transparency, and provide legitimacy to the use of public funds and resources: it provides citizens and other stakeholders with information as to whether public decisions, including in terms of allocating budgetary resources and implementing new laws and regulations, are producing the expected results (OECD, 2018^[1]).

Until recently, very limited comparative information was available in terms of countries' actual practices in the area. This note presents the highlights from a comparative analysis of 42 countries, including 35 OECD countries, and 7 non-members, which represents one of the largest data collection of its kind in the area of policy evaluation. The information was collected through the OECD Public Governance Committee from 2018 until mid-2019, focused around the three pillars of institutionalisation, quality and use which are critical to ensure that policy evaluation systems can make a decisive contribution to sound public governance. More information on the survey and methods is available in the full report (OECD, 2020^[2])

In general, governments are conducting evaluations to attain a significant range of policy objectives as shown by the survey results (Figure 1). According to the survey results, countries show a strong commitment towards policy evaluation, even if there are no clear-cut priorities among these objectives aside from a mainstream concern to measure government's performances/outputs, as well as to promote evidence-informed policymaking, and improve the quality of public services.

Figure 1. Government's main objectives for conducting evaluations



Note: n=42 (35 OECD member countries). Answers reflect responses to the question, "What are the government's main objectives for conducting evaluations?", where 0 indicates "least important objective", 5 is "Neutral", and 10 is a "principal objective".

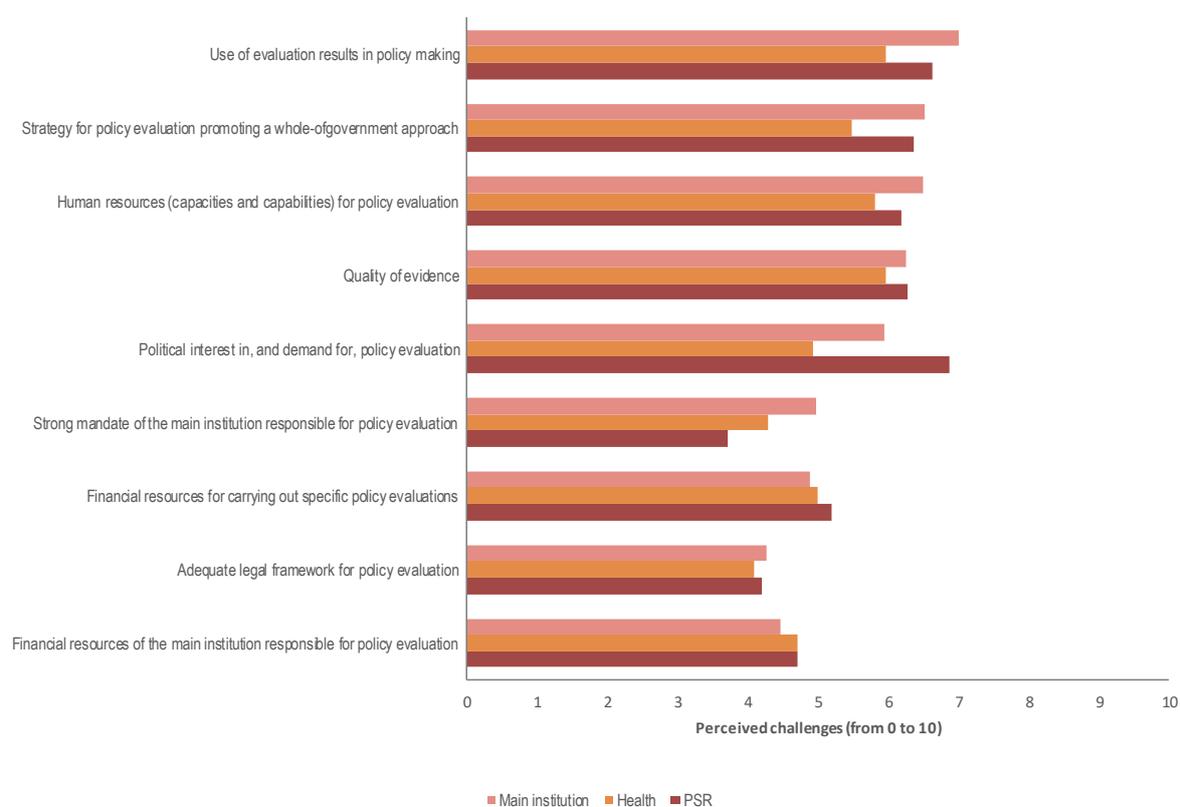
Source: OECD Survey on Policy Evaluation (2018).

However, policy evaluation remains an unfinished agenda in many countries

Despite the growing interest for and acknowledgment of the need to develop a policy evaluation framework, evaluation is often the weakest link in the policy cycle. Countries face technical barriers for carrying out evaluations. These include the challenge of bridging and sharing the knowledge between academic/technical experts and policy makers, who may have different goals, overcoming the heterogeneity and fragmentation of the social and political context and the challenges faced by governments to create and share good, useable and unbiased data within and outside public administration (Rutter, 2012^[3]).

The survey findings reflect the challenges for promoting policy evaluation across government as revealed by responding countries (See Figure 2). These include **the limited use of evaluation results in policymaking**, **the absence of a strategy for policy evaluation** that promotes a whole-of-government approach, and **the limited availability of human resources** (capacities and capabilities). The sectoral surveys, highlighting the perceptions of Ministries of Health, and Ministries of Public Sector Reform, provide additional insights. In the area of health, the quality of the evidence is perceived as a challenge, while Ministries of PSR highlight the political interest in and demand for policy evaluation as a challenge.

Figure 2. Challenges for promoting policy evaluation across government



Note: For the main institution n=42 (35 OECD member countries). For the Health ministry n=31 (28 OECD member countries). 9 countries (7 OECD member countries) did not participate in this survey. Moreover, 2 countries (1 OECD member country) are not included as they answered that none of the policies that fall under their institution's responsibility are evaluated. For the PSR ministries n=25 (20 OECD member countries). 11 countries (10 OECD member countries) did not participate on this survey. Moreover, 6 countries (5 OECD member countries) are not included as they answered that none of the policies that fall under their institution's responsibility are evaluated. Answers reflect responses to the questions, "What are the government's current challenges for promoting policy evaluations?" for the main institution and "What are current challenges for promoting policy evaluation in your institution?" for Health and PSR, where 0 indicates that is a "rare challenge", 5 is "Neutral", and 10 is a "principal challenge".

Source: OECD Survey on Policy Evaluation (2018).

Framing a systemic perspective: promoting institutionalisation, quality and use of policy evaluation

Governments are confronted by complex challenges, as they have to face volatile, uncertain, complex and ambiguous issues, which have called for system approaches to public sector challenges (OECD, 2017^[4]). These issues are interrelated in multiple, hard-to-define ways. Traditional approaches are no longer delivering the expected results, as they reflect poor design and siloed sectoral approaches rather than whole-of-government initiatives. While policy evaluation can be used in a fragmented way, as an activity that is essential to accomplishing distinct tasks, it plays a transversal role in supporting effective policy design and decision-making across governments as a whole. Therefore, the various efforts conducted at different levels of government should be understood from a holistic perspective, as part of a policy evaluation system. A policy evaluation system is one in which "evaluation is a regular part of the life cycle of policies and programmes, is conducted in a rigorous and systematic manner and the results of which are used by decision-makers while making information available to the public" (Lázaro, 2015^[5]).

This study examines evaluation through the **triple lens of institutionalisation, quality and use** from a governance perspective. Attention to these three components will ensure that policy evaluation contributes to sound public governance across the policy cycle, helping to ensure government's policies, programmes and initiatives are achieving their intended goals (OECD, 2019^[6]).

Box 1. Components of the policy evaluation system

According to academic literature, a policy evaluation system rests on three components, which can be defined as follows:

- **Institutionalisation** is defined as the systematic process of embedding evaluation practices into more formal and systematic approaches. It can include establishing an evaluation system in governmental settings through specific policies or strategies (Lázaro, 2015^[5]) (Gaarder and Briceño, 2010^[7]).
- **Quality** is defined as policy evaluations that are technically rigorous as well as well governed; that is independent and appropriate for the decision-making process (Picciotto, 2013^[8]).
- Policy evaluation **use** is defined under three conditions (Ledermann, 2012^[9]):
 - Symbolic use (also known as persuasive), occurs when the results of evaluations are taken up to justify or legitimise a pre-existing position, without changing it;
 - Conceptual use occurs when evaluation results lead to an improved understanding or a change in the conception of the subject of evaluation;
 - Instrumental use corresponds to cases where evaluation recommendations inform decision-making and lead to an alteration in the object of evaluation.

Source: (OECD, 2020^[2])

Despite some achievements, anchoring a fully functional evaluation system remains a work in progress

Countries show mixed practices in regard to their approach to institutionalisation. While in some countries policy evaluations are promoted through a whole-of-government policy framework, where a central institution is given wide responsibilities, in others, evaluations are the sole responsibility of line ministries. In most cases, institutionalisation of policy evaluation relies on a range of different actors, including centres of government, the treasury, ministries of planning, as well as autonomous agencies. These actors are given varying mandates and levels of responsibility. Yet, institutionalisation is not a panacea as in some countries institutionalisation is well developed on paper, but in practice the implementation is still lagging behind.

Countries are investing in a range of tools to promote quality and use. A key component of quality includes investment in public sector skills, particularly policy advisory skills. Some methodological tools are also being developed to provide guidelines on *whether* and *how* government priorities should be evaluated; according to *what criteria*; and *what should be evaluated* (inputs, activities, outputs, outcomes). These guidelines, or standards, serve in many countries as reference guides for the evaluation process. In some cases, countries also implement quality control mechanisms, and the institutions beyond the executive, such as Supreme Audit Institutions and Parliaments usually play a role in this regard, as part of a broader effort to ensure accountability.

Use remains perhaps the greatest challenge for evaluations. There are multiple reasons, which are related to the understanding, the timeliness of evaluations, as well as to how they are communicated within and outside the public sector. Some general mechanisms have been identified to foster use, and it would be useful to reflect on their effectiveness. These include conducting utilisation-focused evaluations, facilitating the communication and dissemination of evaluation results, increasing demand for evidence by promoting competencies for use of evaluations, and creating an evaluation market place by embedding use of evidence in the institutional set-up.

The rich results illustrate the potential for sharing good practices across countries in the future, and identifying what works and what does not. Given that evaluation requires significant investments upstream in the policy making chain, many countries have embedded some guidance in their policy-making processes and their institutional apparatus. Strengthening this guidance by offering broader international insights may help to better anchor evaluation into the institutional and governance frameworks, while respecting the flexibility and diversity of approaches. The short term objectives are to make evaluations happen, and to ensure their relevance, quality and impact. In a sense, the ultimate goal should be to deliver public value, increase the effectiveness of the public sector as a whole, and to provide government, the economy and civil society with a shared understanding of the challenges, the advantages, costs and implications of evaluations.

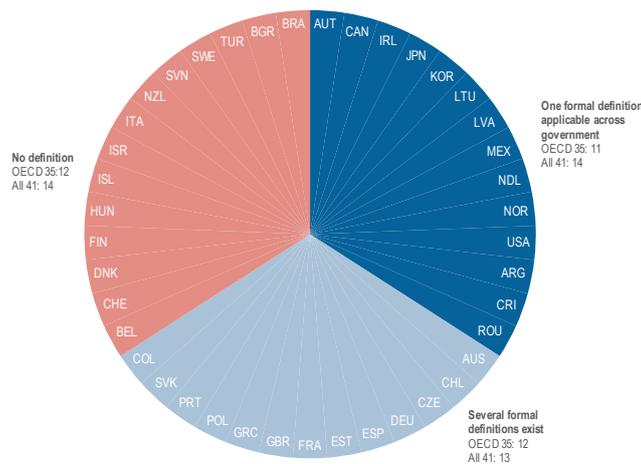
Definitions, methodology and integration into the policy making chain

The role of definitions

Evaluation may be needed to understand why and how a policy was or has the potential to succeed (i.e. learning) and benefits from an efficient allocation of financial resources (i.e. accountability). However, multiple tools and practices are available to measure success and assess whether government actions have met their expected goals. Such tools may include the monitoring of government programmes, spending reviews and performance management – which may complement each other. In this respect, it is important to define and understand the specific role of policy evaluation, across a range of other practices that can measure and promote government’s effectiveness and efficiency.

There is therefore a benefit to ensuring an appropriate definition of policy evaluation. The findings show that more than half of the survey respondents (27 countries) have a definition of policy evaluation: 14 of them have one formal definition applicable across the government, while 13 have several formal definitions and 14 do not have one (Figure 3). In general, countries’ definitions are quite heterogeneous and tend to reflect their own institutional set up.

Figure 3. Government’s formal definition of policy evaluation



Note: n=41 (35 OECD member countries). Kazakhstan answered that they do not know if there is a formal definition for policy evaluation. Answers reflect responses to the question, “Does your government have a formal definition of policy evaluation?” and “Please provide the definition/s and the reference to the relevant documents”.

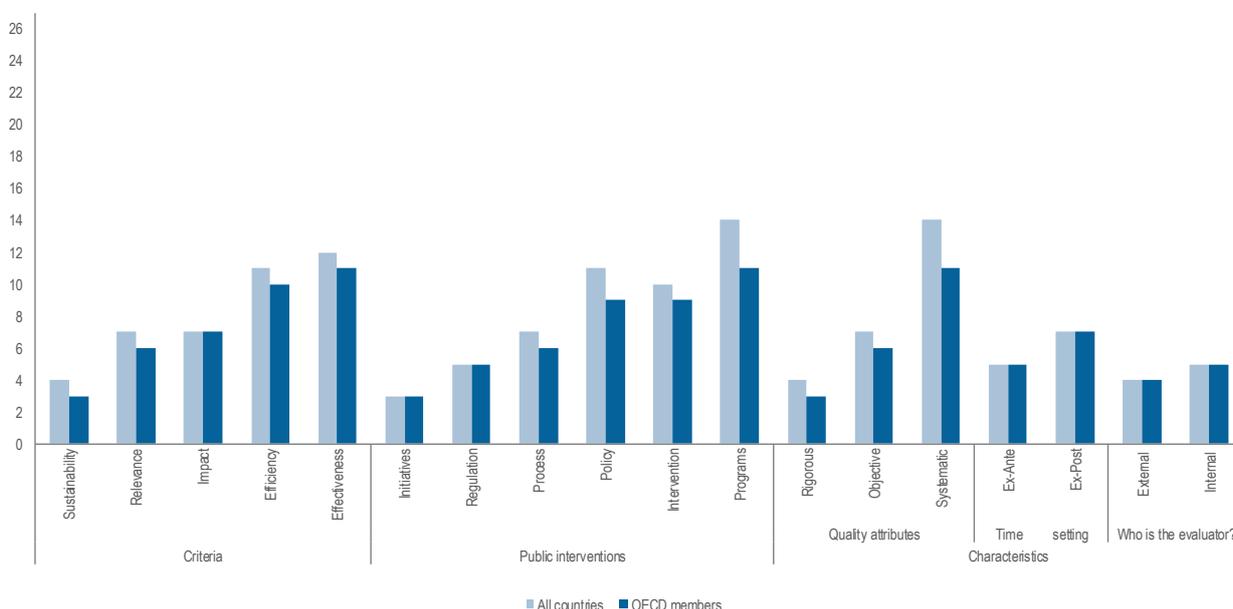
Source: OECD Survey on Policy Evaluation (2018)

Multiple definitions with some common conceptual approaches

Countries have adopted a variety of definitions to approach policy evaluation. This report and the survey sent to countries to collect data, use one such definition, which was used as a starting point to identify common elements underpinning country definitions.

Common elements in country definitions were found, which can be grouped using a clustering approach (Figure 4). These include *what* should be measured (policies, programs, plans, reforms), *why* an evaluation should be conducted (criteria), *when* (ex-ante or ex-post), and *who* is involved. In practice, most of the countries include elements related to the public interventions evaluated, and the assessment criteria as well as quality attributes. However, the elements of time setting, and the type of evaluators involved are less represented.

Figure 4. Conceptual clusters included in the definition of policy evaluation



Note: Answers reflect responses to the question, "Does your government have a formal definition of policy evaluation?" and "Please provide the definition/s and the reference to the relevant documents".

Source: OECD Survey on Policy Evaluation (2018)

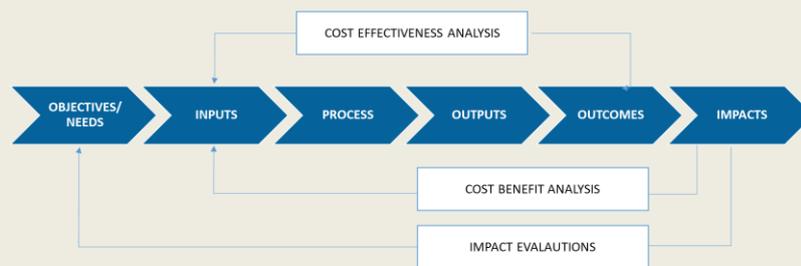
A range of methodologies and tools are used by countries

The findings help to shed light on the degree and nature of policy evaluation definitions across countries. This requires building a shared terminology for some of the different types of evaluations along the policy cycle chain. The elements that are most commonly evaluated in the policy chain include outputs (93%) and outcomes (97%), followed by impact (86%), input (83%), and process (79%). In spite of the potential of Randomised Control Trials (RCTs or QEDs) to provide more rigorous findings, RCTs are only used in 28% of the sample. Moreover, methods such as regressions/ econometrics/ structural equation modelling are used by 59% of the sample for the evaluation of government wide priorities, by 39% of the sample in the case of the health related aspects and by 32% of the sample for the respondents in the area of public sector reform.

Box 2. The Policy chain cycle

The figure below highlights the policy chain and describes the interaction between the **needs** or problems identified in a specific context, the corresponding intervention to solve them (**inputs, process and outputs**), its **outcomes**, and evaluation (**impact**); specifically, each element is defined as follows (OECD, 2016_[10]):

1. The **input level** refers to the resources employed to implement a policy such as staff, money, time, equipment, etc.
2. The **process level** refers to the activities that were undertaken in a policy.
3. The **output level** refers to a first level of results, directly associated with the products delivered by the policy implemented.
4. The **outcome level** refers to the medium-term (directly) consequences of the policy implemented.
5. The **impact level** looks at the long-term consequence of a policy initiative ,



In practice, things are often more complex, given the scope of an evaluation: evaluating a single initiative, versus a comprehensive action plan, requires different tools and can probably reach different levels of understanding (OECD, 2016_[10]). Different methods can be used to measure impact such as efficiency (Cost effectiveness analysis) or effectiveness (Impact evaluations). For instance, a Cost Effectiveness Analysis would required the comparison between cost (from the inputs or resources employed) and the outcomes obtained.

Sources: OECD (2016_[10]).

Why is institutionalisation so difficult to achieve?

A sound institutional framework constitutes an important building block for implementing a functioning evaluation system. Institutionalisation can contribute to providing incentives for the systematic conduct of evaluations, and clear rules for transparent and clear management. However, countries still face major challenges in the institutionalisation process, particularly in achieving consistent frameworks from a whole of government perspective. The availability of human resources, in terms of capacity and capabilities also represents major gaps. Finally, sustaining political interest and demand for policy evaluation remains challenging.

Sound institutional frameworks can align isolated and unplanned programme evaluation efforts; improve the comparability of results across time, institutions, and disciplines; and promote transparency and accountability in the management of evaluations. Sound institutional frameworks therefore play an important role in protecting policy evaluation practices from undue political influence and misplaced bureaucratic practices (Gaarder and Briceño, 2010^[7]) (Parkhurst, 2017^[11]) (Lázaro, 2015^[5]). Nevertheless, the findings show that countries still face important challenges in this area (Figure 5):

- The whole-of-government strategy is a significant challenge for all countries. Institutionalising evaluation involves multiple actors, which usually causes resistance to change. Hence, without effective guidance, organisations may fail to make coordinated decisions and agree on a common vision, mission and goals (Cinar, Trott and Simms, 2018^[12])
- The human resources are another lacking area. This is partly related to the fact that civil servants lack the time to absorb the new practices associated with evaluation, especially when those are not directly related to their operational priorities.
- Finally, ensuring political interest and demand for policy evaluation is also an issue. In this regard, finding incentives for civil servants and experts in charge of advocacy to engage in tasks outside their immediate area of responsibility is an issue reflected by a number of respondents.

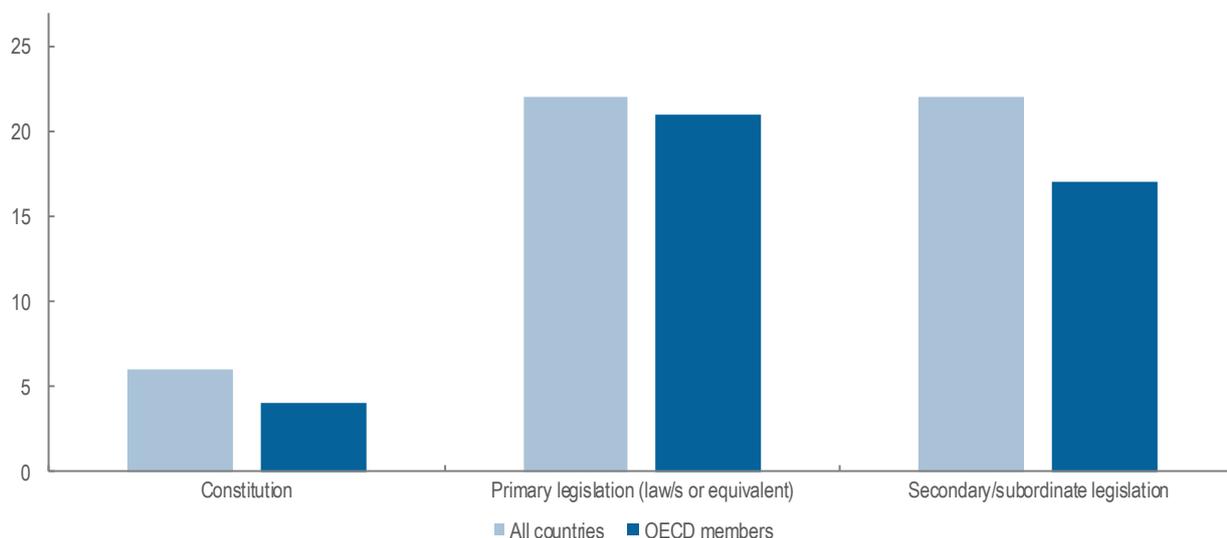
Evaluation is generally well embedded in legal and regulatory frameworks

An adequate legal framework can provide a clear mandate and guidance to institutions. Around two-thirds of responding countries have created a legal basis for policy evaluation. This demonstrates the importance that countries attribute to the systematisation of this practice across government.

Legal frameworks outline specific paths for government-led institutionalisation. A majority of countries (29 countries, 23 OECD countries) have developed a legal framework (Figure 5). However, the legal framework can vary substantially across countries. The most frequent legal basis for institutionalisation is

primary (22 countries) and secondary legislation such as decrees and ministerial resolutions (22 countries). Six countries including Germany, France, Mexico, and Switzerland have anchored policy evaluation as a constitutional duty.

Figure 5. Legal frameworks guiding policy evaluation across government



Source: OECD Survey on Policy Evaluation

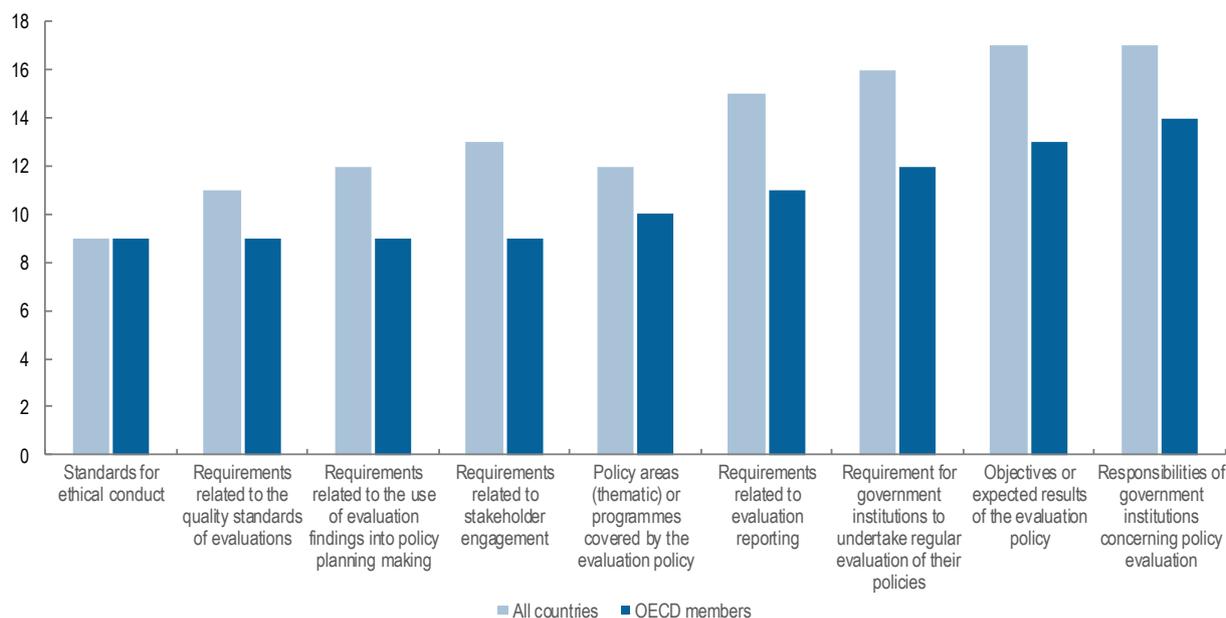
Note: n=28 (22 OECD member countries). 14 countries (13 OECD members) are not included as they answered that they do not have a legal framework guiding policy evaluation across the government. Answers reflect responses to the question, “At what level is policy evaluation across government legally embedded? (Check all that apply)”.

The role of a supportive overall Policy Framework

Besides legal frameworks, which stipulate regulatory requirements, overall policy frameworks have an important role to play, as they can explicitly state the role of evaluation and its importance for aligning goals with policy, which is essential for pursuing an evaluation culture. Half of the countries in the sample have developed such a policy framework for organising policy evaluation across government, aside from or as part of the legal framework (Figure 6). Among those, around 19 countries implemented both a legal and a policy framework, including 15 OECD countries.

Policy frameworks most commonly include an outline of government institutions’ responsibilities concerning evaluations, as is the case with a total of 17 countries surveyed (of which 14 are OECD) such as Greece, Japan, Korea, and Lithuania. Evaluation plans, or requirements for government institutions to undertake regular evaluation of their policies are also common among countries, for instance in the cases of Latvia, Mexico, and Poland (16 overall, and 12 OECD countries have these). Other frameworks state objectives or expected results for the evaluation policy (17 countries overall including 13 OECD ones), such as Slovakia and Spain. Additionally, policy frameworks may focus on particular policy areas and programmes, as is the case for 12 countries surveyed, including 10 OECD countries who have implemented thematic or specific policy frameworks, such the case of Argentina and Costa Rica.

Interestingly, some countries who lack an overarching legal framework have nevertheless created a policy framework to promote evaluations and their institutionalisation. These include Canada, the United Kingdom and Ireland, perhaps due to some form of a Westminster tradition, and implicit unwritten rules.

Figure 6. Availability and content of Policy Frameworks

Note: n=19 (15 OECD member countries). 21 countries (18 OECD member countries) are not included as they answered that they do not have a policy framework for organizing policy evaluation across government. Data is not available for Canada and Ireland. Answers reflect responses to the question, "Which element(s) do(es) the document/s referred to under Q4 and Q5 cover concerning policy evaluation across government? (Check all that apply)". The documents referred to under Q4 and Q5 are the ones stipulating a policy framework organising policy evaluation across government.

Source: OECD Survey on Policy Evaluation (2018).

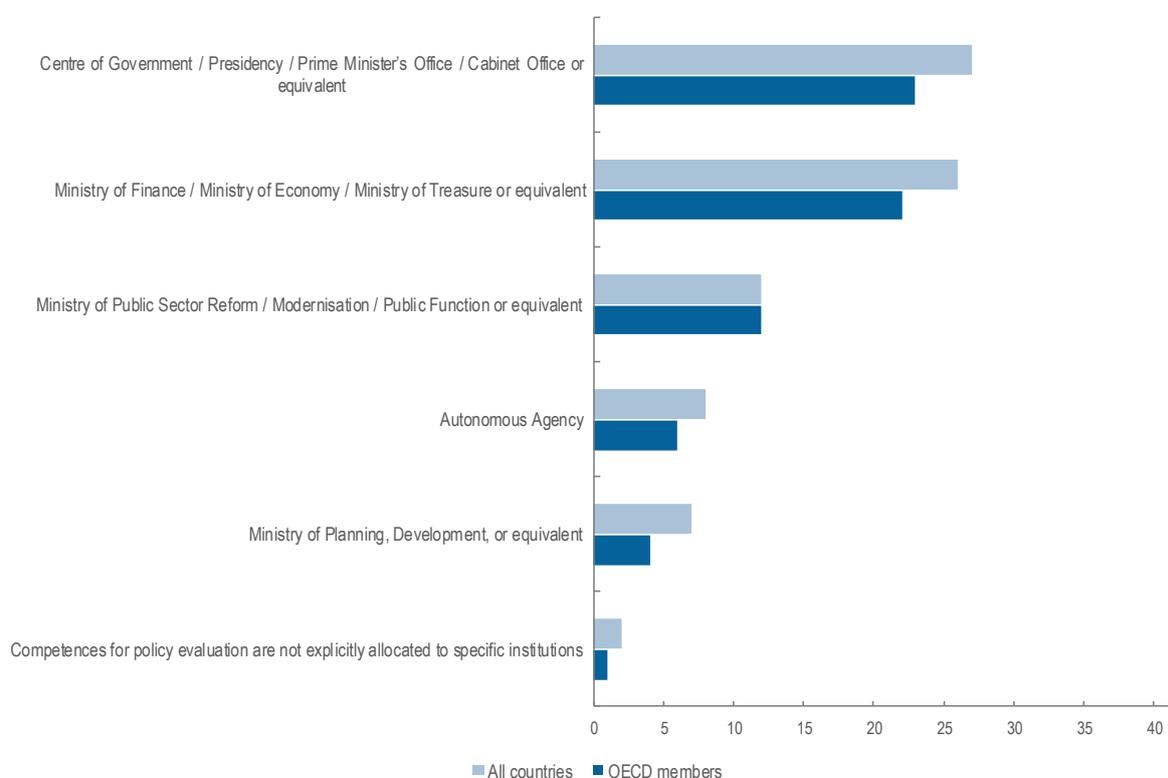
The significant role of policy evaluation guidelines

Guidelines can assist policy makers in conducting policy evaluation successfully. They intend to assist all those participating in the implementation of a policy in better planning, commissioning and managing its evaluation (Austrian Development Cooperation, 2009^[13]), and considering risks and implementation issues (Department of the Prime Minister and Cabinet (Australia), 2013^[14]). A full presentation of the guidelines identified in the analysis is available in OECD (2020) and will be the matter of further investigation. A majority of countries (31 countries and 26 OECD members) are equipped with such guidelines to support the implementation of policy evaluation across government. Nonetheless, similar to the case of policy frameworks, some countries only have guidelines but do not report a policy or a legal framework, such as Australia, Finland, New Zealand and Portugal.

The key institutions in charge of policy evaluation and their mandates

Almost all countries (40 of 42 participating countries, including 34 OECD countries) allocate responsibilities for policy evaluation across government. A great majority of countries have chosen to allocate the mandate of coordinating policy evaluation across the executive to either one or several institutions often close to a locus of power, either at the Centre of Government, or at the Ministry of Finance or other ministries endowed with significant cross-departmental coordination capacity. The fact that in the majority of OECD countries more than one institution conducts policy evaluation underlines the importance and necessity for strong steering and coordination capacities.

Figure 7. Institutions within the Executive that have competences related to policy evaluation across government



Note: OECD Survey on Policy Evaluation (2018).

Source: n=42 (35 OECD member countries). Answers reflect responses to the question, "Which of the following institutions within the executive have competences related to policy evaluation across government? (Check all that apply)".

To decide where to locate the key function of driving policy evaluation, countries need to weigh up the trade-offs between political influence and proximity to decision making power such as for the CoG or the Ministry of Finance on the one hand, and managerial independence and accountability provided by an autonomous agency, on the other. While institutionalisation choices are often determined by countries' history and political traditions, identifying broader patterns of institutionalisation may require further analytical work in the future.

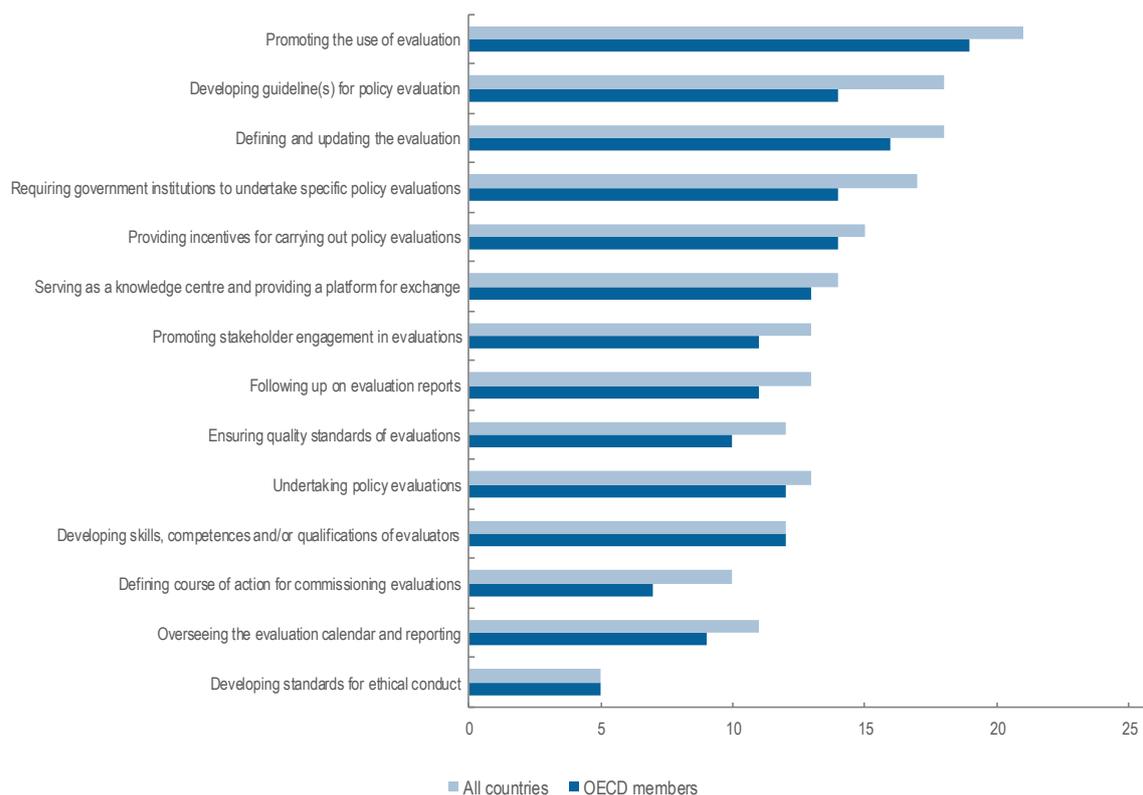
Centres of government have the widest responsibilities

The Centres of Government (CoG) play a crucial role in embedding a whole-of-government approach to policy evaluation. An assessment of the mandate of countries' CoG reflects its role as a guiding institution in policy evaluation across government.

The Centre of Government (CoG) is where policy evaluation is most often institutionalised. From those countries that assigned responsibilities for government wide policy evaluation to the CoG, nineteen have tasked it with promoting the findings' use across government. The CoG as an institution close to political decision-making is well placed to ensure that the evaluation results are fed back into the policy-cycle. The twin roles of the CoG and the ministries of finance are reflected through the OECD Survey data. Accordingly, of the 26 countries (22 OECD countries) which noted that the Ministry of Finance or equivalent has a role in policy evaluation across government, 19 countries (16 OECD countries) also mentioned the CoG. However, aligning the role of the CoG and the Ministry of Finance may require some internal

coordination and may represent a challenge to focus and align government action to achieve shared outcomes (GOV/PGC (2015)3). According to the OECD report on the CoG (2017), policy coordination and monitoring the implementation of government policy are among the five key responsibilities of the COG. The COG has a crucial role to play to ensure that decision makers are able to debate the merits of policy options based on the evidence-informed policy advice presented in a way that helps understand the costs, benefits, likely impacts and implications of those options.

Figure 8. Mandate of the Centre of Government for Policy Evaluation



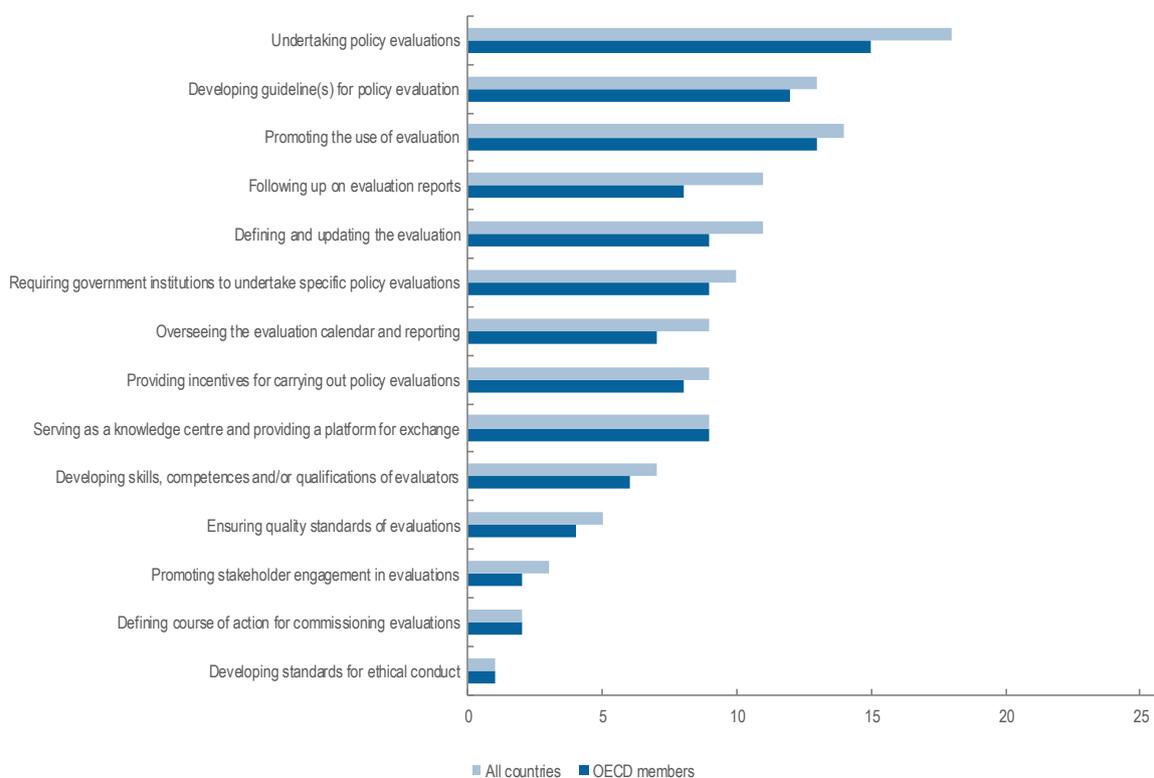
Source: OECD Survey on Policy Evaluation (2018)

Note: n=27 (23 OECD member countries). Answers reflect responses to the question, "Please list the duties and responsibilities of this/these institution/s related to policy evaluation across government" for the Centre of Government / Presidency / Prime Minister's Office / Cabinet Office or equivalent. Answer option "Other" is not included.

The roles of ministries of finance

In a number of countries, Ministries of finance play a key role for the institutionalisation of policy evaluation as they are charged with providing incentives to enhance the quality of public expenditures and improve the results of government's spending, including through improving performance management and developing spending reviews. In addition, policy evaluation also has a role to play to increase the quality of major future public investments, including in the area of infrastructure, which is another area frequently under the remit of Ministries of finance.

Figure 9. Mandate of the Ministry of Finance/Economy/Treasury or equivalent for policy evaluation



Note: n=26 (22 OECD member countries). Answers reflect responses to the question, "Please list the duties and responsibilities of this/these institution/s related to policy evaluation across government" for the Ministry of Finance / Ministry of Economy / Ministry of Treasure or equivalent. Answer option "Other" is not included.

Source: OECD Survey on Policy Evaluation (2018)

The results show that Ministries of Finance are the second most frequent institution (26 countries in total, including 22 OECD countries) with competences for policy evaluation across government. The ministries' mandate reflect the ministries' strong role as coordinating institution that provides guidance and in some cases supervision to other ministries' evaluation activities. Eventually, similarly to the CoG, the ministry of Finance's mandate includes the promotion of the use of evaluation findings. These finding can be embedded in the ministries of finance traditional task of ensuring value-for-money. From a budgeting perspective, evaluation is part of a "Performance Continuum", which requires an integrated approach to evaluation, spending reviews and performance management. However, such an approach implies that ministries and the administration be equipped with the right skills and capacity to effectively measure outcomes and outputs and gather the data for evaluation.

Other actors within the executive: ministries of planning and agencies

Other institutions such as the Ministry for Planning and Development as well as autonomous agencies can yield a significant contribution to policy evaluation. This is true in a smaller share of the sample where Ministries of Planning play a significant role (7 countries, and 4 OECD countries: Chile, Czech Republic, Poland and Slovakia). This can be traced back to the strong role that national development plans have in the region and the necessity to build capacity to evaluate their implementation. These strategic plans tend to be evaluated by the ministries of planning or development, which gives them with a strong mandate and experience for policy evaluation across government.

In addition to institutions within the executive, autonomous agencies in the OECD have taken up competences related to policy evaluation across government in Mexico, Poland, Spain, Sweden and the United States, thereby further diversifying the institutionalisation approaches of policy evaluation. These autonomous agencies tend to be characterised by stronger accountability mechanisms due to their budgetary independence. Agencies play a significant role in a Nordic setting, such as Sweden or Norway, even if the peculiarities of the survey may limit such identification in the results.

An important function of Supreme Audit Institutions and a role for Parliamentary budget offices

While the current findings were focused more on institutional arrangements within the executive, a full understanding of policy evaluation, as a tool for broad political accountability, requires understanding the role of institutions beyond the executive. Generally, Supreme Audit Institutions (SAIs) are tasked to conduct ex post audits of policies and programmes. SAIs are moreover in charge of evaluating the efficiency and effectiveness of the government's policies and function as crucial actor of transparency and accountability. In the case of OECD countries, 27 Supreme Audit Institutions (SAIs) (33 in all countries) received this mandate, according to the survey. There is also a group of Supreme Audit Institutions, under INTOSAI, which federates experts from supreme audit institutions interested in Policy Evaluation, which has also produced some cross-country results available in the main report.

In addition, Parliaments or Congress, and Parliamentary budget offices can play a role: Around one third of responding OECD countries (11) and (13 countries in total) involve either Congress or the Parliamentary Budget Office in policy evaluation to some extent, whereas none of the aforementioned institutions has a mandate on policy evaluation in 7 OECD countries (8 countries in total).

Fostering an evaluation culture through the promotion of quality and use of evaluation

The importance of quality and use

A positive evaluation culture is an essential component of a functioning policy evaluation system. This requires framework conditions so that evaluations of a decent quality can be produced and that they can be fed in effective ways into policymaking processes. Sustaining a robust evaluation system requires more than the institutionalisation of policy evaluation: evaluations should not only be conducted, they should be of high quality and used. Quality evaluations contribute to better decision-making, policy learning, and accountability. Conversely, while high quality evaluations may be produced, they may fail to be integrated in policy decisions. Moreover, quality and use are inherently interrelated: quality makes evaluations credible and robust, meaning that they can soundly inform decision-making and be confidently used by decision-makers, stakeholders and citizens. Using quality evaluations also strengthens decision-makers' accountability, as it ensures that they provide cost-efficient and quality services to citizens. Fostering both is therefore essential to embedding policy evaluation throughout the policy cycle.

What are quality evaluations?

The quality of policy evaluation covers two key dimensions:

- *Quality of content*, that policy evaluations are technically rigorous across a range of methodological standards.
- *Quality of the Governance process*. This implies that the process for conducting the policy evaluation and feeding the results into policy-making is well governed, and presents guarantees in terms of independence of analysis and appropriate mechanisms for transparency and external engagement, thus reducing the scope for undue influence.

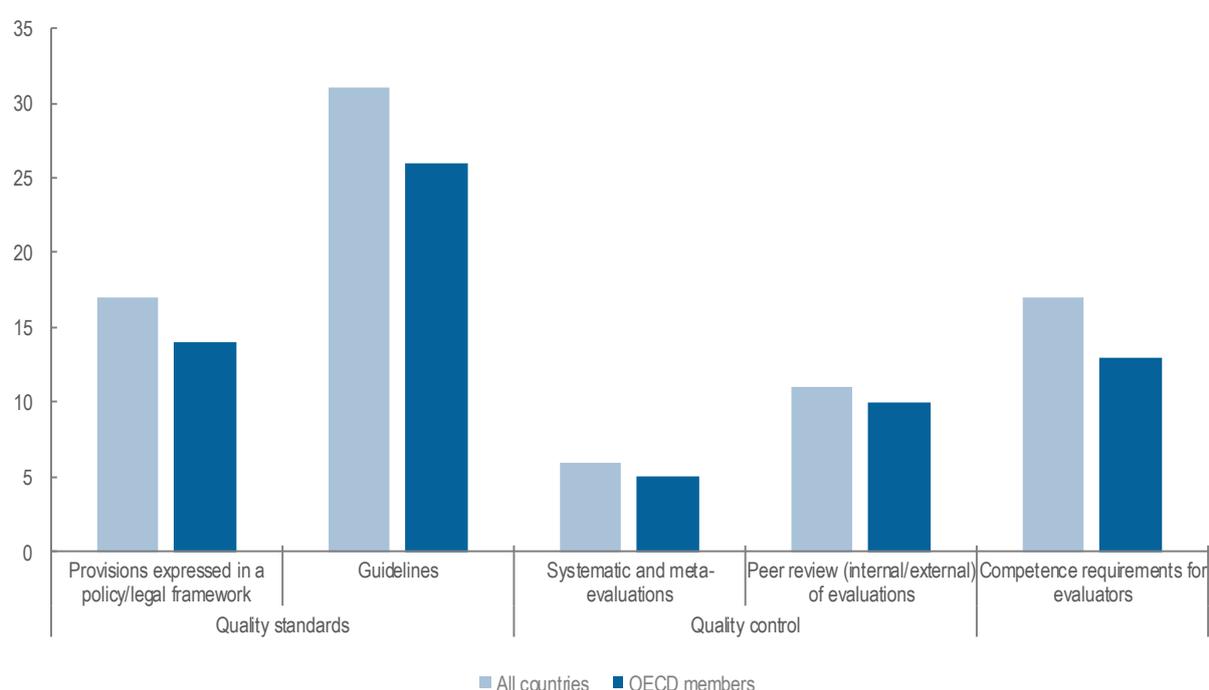
Technical quality can be achieved through proper design, sound data collection, rigorous methods and adequate resourcing for evaluations. Nevertheless, technical quality is necessary but not sufficient to promote an evidence informed approach to policymaking, particularly one that will increase trust in the findings and effectively promote accountability vis-a-vis citizens and stakeholders. Evidence informed decision-making and the use of evidence remain inherently political, as evaluations are but one source of information in a messy political process that also reflects diverging values, interests and opinions, and can be subject to bias and influence. Still, there are ways to ensure that the evaluation process is well governed, that it takes into account the need for integrity, equity, and reflects different values, for instance by ensuring proper engagement with stakeholders and weighing different political considerations.

How can countries promote the quality of the evaluations?

Countries ensure evaluation quality through four key mechanisms: quality assurance, quality control, competences and institutions outside the executive. The survey results show that countries have sought to promote the quality of evaluations via three main channels (Figure 10):

- developing standards on the quality of the evaluation process, which can be embedded in evaluation guidelines or in legal/policy frameworks;
- controlling the quality of the evaluation end product, either through systematic or meta evaluation, or peer review;
- supporting and promoting evaluator competences.

Figure 10. How do countries ensure the quality of evaluations across government?



Note: n=42 (35 OECD member countries). Answers reflect responses to the questions “How does your government ensure the quality of evaluations across government” and “Are there guidelines available to support the implementation of policy evaluation across government?” Systematic and meta-evaluations refer to the evaluations designed to aggregate findings from a series of evaluations.

Source: OECD Survey on Policy Evaluation (2018).

Assuring quality through standards for policy evaluation

Countries have developed quality standards that enable ensuring a degree of consistency in the design and process of evaluations, thus ensuring that the evaluation process respects certain quality criteria. These standards are usually embedded in evaluation guidelines and, less frequently, in legal or policy frameworks. The survey information shows that most countries have developed standards regarding both the technical quality of evaluation and its good governance, reflecting their understanding of the dual determinants of quality evaluations.

A majority of countries have developed guidelines that seek to address both the technical quality of evaluations and the good governance of evaluations. In Estonia, for instance, the Methodology of Impact

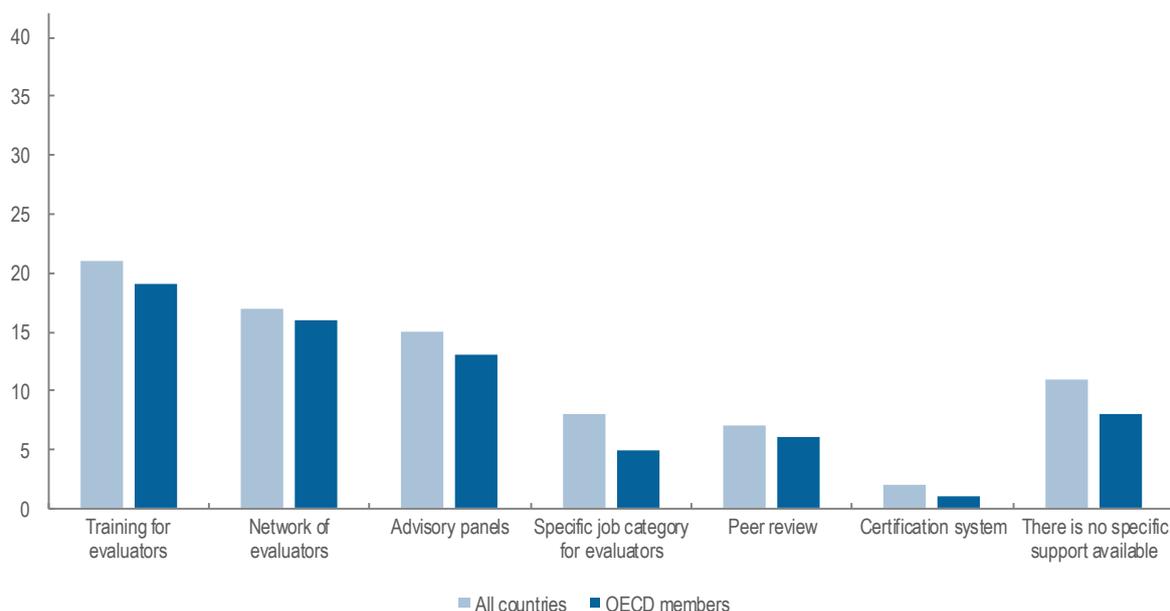
Assessment (2012) guidelines describe the technical features of impact evaluations of policies and programmes, while the Good Public Engagement Code of Practice (2012) focuses on the principles for the good governance of evaluations, such as the involvement of the public and interest groups in decision-making processes. There are therefore close interactions between the quality of evaluation and the opportunity for engaging with citizens and ensuring open government practices.

As a result, a number of the standards for use are related either to budgeting frameworks, and public spending codes, or to mechanisms to ensure the quality of impact assessment or regulatory impact assessment. These are well developed in the regulatory policy area, where they are also reflected in the [OECD Recommendation of the Council on Regulatory Policy and Governance](#).

Controlling the quality of the evaluation product

Mechanisms for quality control ensure that the evaluation design, its planning and delivery are meeting certain quality criteria. While quality assurance mechanisms seek to ensure credibility in how the evaluation is conducted (the process), quality control tools ensure that the end product of the evaluation (the report) meets certain quality standards. Both are key elements to ensuring the robustness of policy evaluations. Mechanisms for quality control implemented by countries include peer review of evaluation products, meta-evaluations and self-evaluation tools and checklists. The findings of the survey (Figure 10) show that such quality control mechanisms are less common than quality assurance mechanisms, with only approximately one third of countries using a quality control mechanism in one form or another.

Figure 11. Building competencies for policy evaluation



Note: For the main institution n=42 (35 OECD member countries). Answers reflect responses to the question “How does your government ensure the quality of evaluations across government”..

Source: OECD Survey on Policy Evaluation (2018).

Building competencies is another way to ensure quality, as individuals who possess the right competences are more likely to produce high quality and utilisation-focused evaluations. Investing in evaluators' competencies implies having the appropriate skills, knowledge, experience and abilities. Countries have generally recognised the crucial role of competencies in promoting quality evaluations. In fact, survey data shows (Figure 11) that a majority of respondents use mechanisms to support the development of evaluators' competences.

These are promoted by training internal and external evaluators, establishing a specific job category for evaluators in government, such as is the case in the US and in Ireland for example,¹ or implementing a certification systems for evaluators. Competency development can be directed at organisations in their entirety, such as through advisory panels and committees, and establishing or supporting a network of evaluators. Detailed results are available in the background report.

The role of Supreme Audit Institutions in promoting quality

The survey results show that Supreme Audit Institutions (SAIs) have taken on an active part in the promotion of evaluation quality. SAIs are key players in the national discourse concerning evaluation quality, usually within countries with a more mature evaluation culture. Thanks to their particular expertise in performance auditing, they give countries external insights on how to improve the quality of their evaluation systems. Supreme Audit Institutions provide useful guidelines for evaluation quality. Lastly, SAIs sometimes perform evaluations themselves, including evaluations of policy evaluation systems, with their own standards for quality.

Tackling the challenge of using policy evaluation

Use of evaluations remains a constant challenge and often falls below expectations. Despite the potential for policies to be based on evidence, an effective connection between evaluation results and policymaking remains elusive. For example, US estimates show that, under the two Obama administrations only 1% of government funding was informed by evidence ((Bridgeland and Orszag, 2013_[15])).

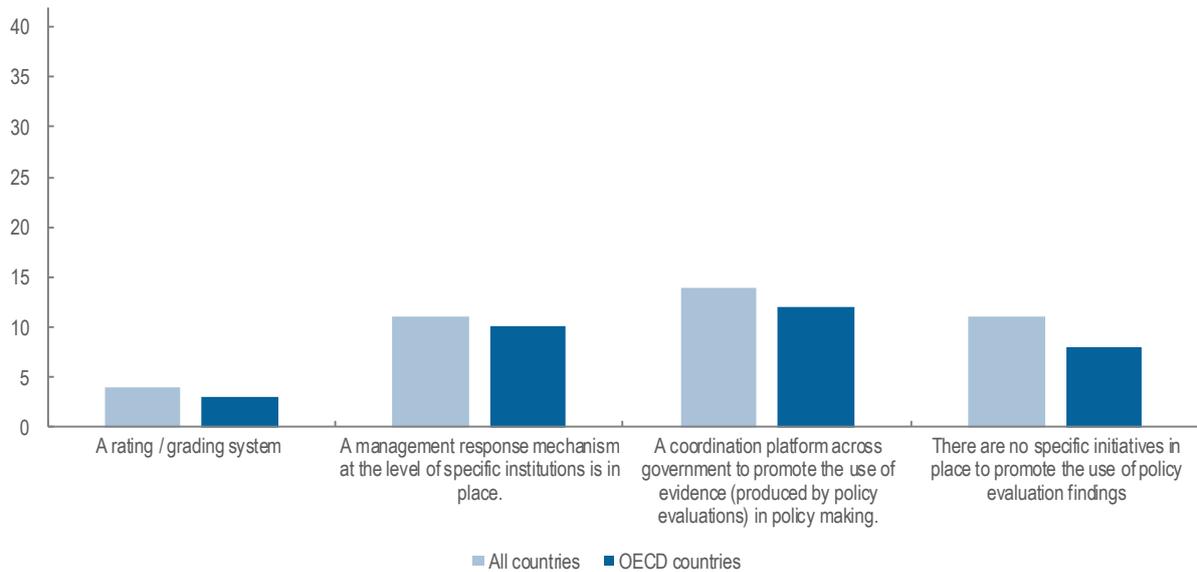
With insufficient use, the rationale for conducting evaluations in the first place is weakened, which may jeopardise future evaluations. One of the most fundamental rationale for conducting policy evaluation is its usefulness in informing policy and decision-making, and improving the quality of public interventions. Evaluation has to produce useful insights on public issues and yield evidence on the impact of policies and their underlying change mechanisms.

In theory, it is possible to promote the use of evaluations through: conducting utilisation-focused evaluative processes; promoting access to evaluations; supporting the uptake of evaluations results; increasing demand for evaluations through competency development; embedding use in the institutional set-up, within and outside of the executive.

Figure 12 provides aggregated survey findings on the broad patterns for mechanisms to promote use, including a grading system, management response mechanisms and coordination platforms, which had been identified as options through the survey.

¹ See OECD (2020)a

Figure 12. How do countries promote the use of evaluation?



Note: n=42. Answers reflect responses to the question “How does your government promote the use of the findings of policy evaluations”. The option “Other” is not included. A rating/grading system is used to classify the robustness of evidence provided and recommendations derived from the policy evaluations exists.

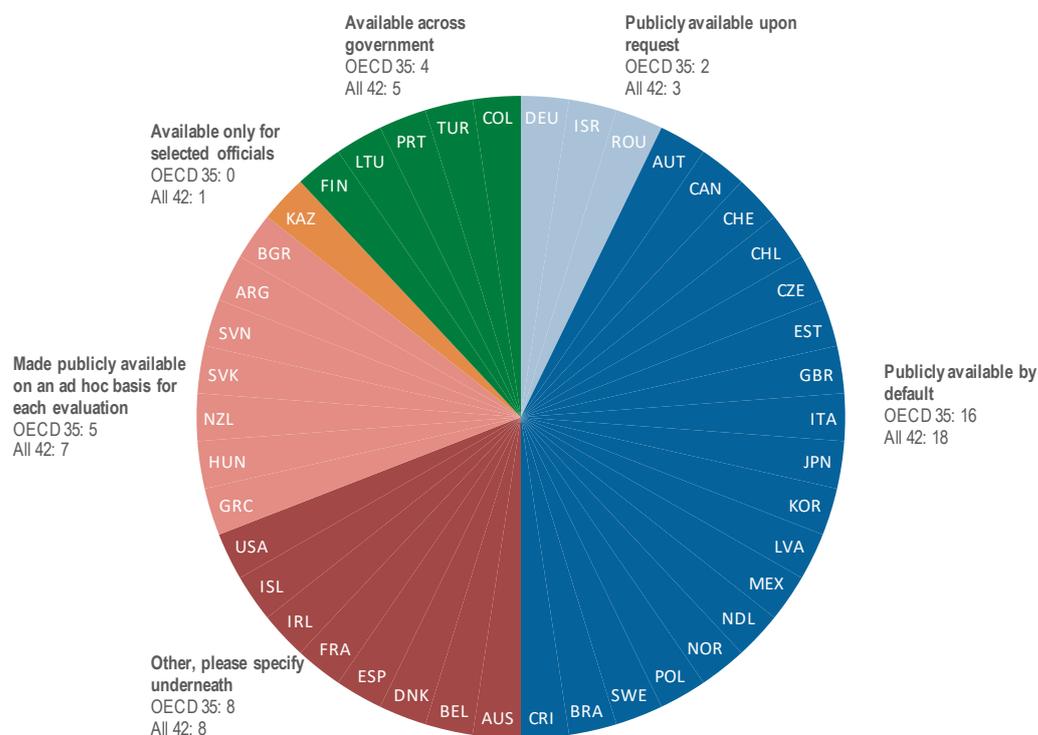
Source: OECD Survey on Policy Evaluation (2018)

Engaging stakeholders in the evaluative process

Conducting *utilisation-focused evaluations* entails engaging stakeholders at every stage of the evaluative process, enabling the evaluation to be fit for purpose and responsive to the needs of their users. Overall, governments are increasingly eager to engage a wide range of internal and external stakeholders in the decision-making process to level the playing field and generate a broader consensus while legitimising more public-policy decisions. There is a consensus that engaging those concerned and affected by evaluations is fundamental to improving design, relevance, transparency and, *in fine*, use. Therefore, stakeholder participation and interaction in the evaluative process can help build trusted relationships and increase the opportunities for evaluation results to impact policy making.

In the Netherlands, for example, the Ministry of Finance’s Regulations for periodic evaluation research, lay down rules for the participation of stakeholders in periodic evaluations (15 March 2018). With each policy evaluation, at least one independent expert must give an opinion on the quality of the evaluation. Similarly, the European Commission’s Better Regulation Guidelines contain a chapter that describes standards for stakeholder engagement. According to these guidelines, views from stakeholders should be included in the evaluation of all programmes and policies issued by the Commission, as well as initiatives with impact assessments.

Figure 13. Making the results of evaluation public



Note: n=42. Answers reflect responses to the question “The results of the evaluation are”. In “Other”, the majority of countries agreed that the public availability of evaluation results depends on the specific agency that commissioned the evaluation, and on its organization.

Source: OECD Survey on Policy Evaluation (2018)

Similarly, communicating findings to stakeholders as the evaluation progresses can favour their adherence to and understanding of the results. Publishing evaluation results is becoming increasingly open and transparent in surveyed countries. Many countries make evaluation findings and recommendations available to the general public by default, for example by publishing the reports on the commissioning institutions’ website (Figure 13). Only one country reported that evaluation results are only available for selected officials on an *ad hoc* basis. Such availability is important for evaluation use, because if citizens are aware of the results and wary of the implications, it may also put pressure on the policy makers to pay attention to the results and ensure that they feed into policymaking. OECD data also shows that uptake of evaluation results by policy and decision makers may be likely when information is easily accessible.

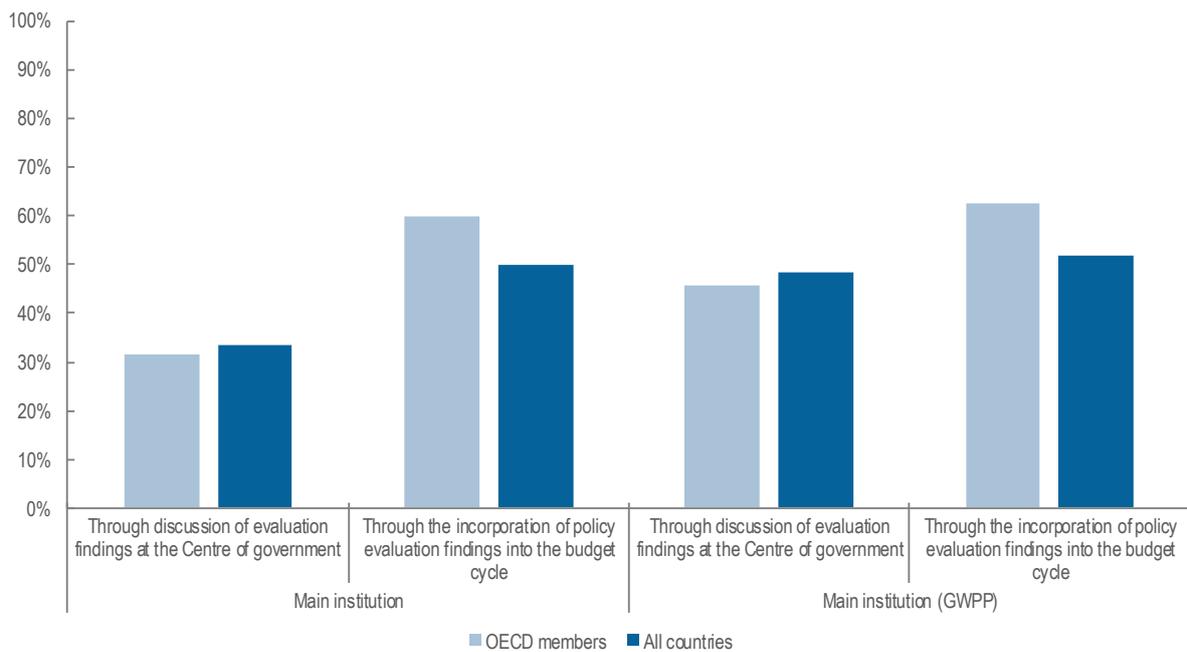
Feeding results into the budget cycle

Incorporation of evaluation findings in the budget cycle is one of the most commonly used mechanism for the promotion of use of evaluations. In fact, OECD data shows that half of surveyed countries report that they incorporate evaluation evidence into the budgetary cycle. Evaluative evidence may be used more or less systematically in the budget cycle. For instance, the OECD identifies three main models of performance budgeting, which reflect the different strength of the links between performance evidence and budgeting (OECD, 2019_[16]):

- presentational (evidence is presented separately from the main budget document),
- performance informed (performance evidence is included within the budget document that is presented on the basis of programmes),
- managerial, direct performance budgeting (direct link between results and resources).

In most OECD countries, performance evidence is included in the budget cycle according to one of the first three approaches (Figure 14). France offers an example of strong links between key performance indicators established at the national level, their *ex post* evaluation, and the budget cycle.

Figure 14. How are policy evaluation findings used into the budget cycle?



Note: The scale is expressed as a percentage of the number of countries concerned, as the number of respondents differs for the main institutions in general as opposed to the main institution for government wide policy priorities. For the main institution n=42 (35 OECD member countries). For the main institution on government-wide policy priorities n=29 (24 OECD member countries). 4 countries (all OECD member countries) answered that they do not have government-wide policy priorities. Moreover, 9 countries (7 OECD member countries) answered that they do not evaluate their government-wide policy priorities. Answers reflect responses to the question "How does your government promote the use of the findings of policy evaluations" and "How does your institution promote the use of the findings of policy evaluations? (Check all that apply)" focused on "Incorporation of policy evaluation findings into the budget cycle".
 Source: OECD Survey on Policy Evaluation (2018).

Discussing evaluation findings at the highest level of government and in Parliament

Countries have also used various mechanisms to discuss evaluation results at the highest political level, such as at the Council of Ministers. This practice is more frequent regarding the evaluation of government-wide policy priorities (GWPP), with about half of countries discussing evaluation findings at the level of the Council of Ministers (or equivalent) in regards to GWPP, against a third of countries for policy evaluations in general. In Korea, for instance, in the context of the “100 Policy Tasks” five-year Plan, evaluation results are discussed at the Council of Ministers.

In addition to the Centre of Government, discussions of evaluation findings also occur at the level of Parliament. For instance, Parliaments have been instrumental in increasing evaluation use by: promoting the use of evaluative evidence in the budgetary cycle by requiring more performance data on government spending, introducing evaluation clauses into laws and commissioning evaluations at the committee level in the context of hearings.

Implications for the future

The findings show overall the key role of policy evaluation as a core element that contributes to sound policy development, strategic planning and performance management through evidence-informed practices. While evaluations are mobilised in various fields, such as for assessing regulatory impacts ex ante or ex post, or to evaluate the effectiveness of public interventions and policies from an expenditure point of view, very few countries have adopted a holistic approach in this area. While many bright spots exist, many countries are also currently facing the digital challenge, which increases the supply and diversity of administrative data that can be mobilised for evaluation purposes.

Yet, at the same time, in a fragmented and diverse political and social environment, maintaining public trust through exemplary accountability processes presents governments with many challenges. Given the significant advances in many countries in this area in recent years, there is scope for taking stock of these findings with a view towards identifying best practices and sharing insights on effective ways to promote trustworthy evidence and evaluation in the policy arena. This is a challenge that many policy makers and public sector analysts are currently tackling across OECD countries and where the opportunities for mutual learning and fast tracking of best practices seem to be highest.

While this document was prepared before the COVID 19 crisis began to worsen among the participating countries, many countries have experienced significant economic and social disruptions since then, calling for sudden government intervention on an unprecedented scale. This will require scaling up the practices for policy evaluation to maintain trust, and ensure that governments can chart a way out of the crisis, which, like the 2008 global financial crisis, may take significant energy and efforts. More than ever, countries will need to improve the effectiveness of public expenditure to restore fiscal space in the longer term, and will need to leverage policy evaluation for this purpose. While it is too early to assess and identify the full legacy of this crisis, many countries will need help in the future to advance on this important and unfinished evidence and evaluation agenda from a public governance perspective.

References

- Austrian Development Cooperation (2009), "Guidelines for Project and Programme Evaluations", [13]
https://www.entwicklung.at/fileadmin/user_upload/Dokumente/Projektentwicklung/Englisch/Guidelines_for_Project_and_Programme_Evaluations.PDF (accessed on 20 September 2019).
- Bridgeland, J. and P. Orszag (2013), *Can Government Play Moneyball? - The Atlantic*, [15]
<https://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/> (accessed on 6 December 2018).
- Cinar, E., P. Trott and C. Simms (2018), "Public Management Review A systematic review of barriers to public sector innovation process", [12]
<http://dx.doi.org/10.1080/14719037.2018.1473477>.
- Department of the Prime Minister and Cabinet (Australia) (2013), *Policy Implementation* [, [14]
<https://www.pmc.gov.au/government/policy-implementation> (accessed on 20 September 2019).
- Gaarder, M. and B. Briceño (2010), "Institutionalisation of government evaluation: balancing trade-offs", *Journal of Development Effectiveness*, Vol. 2/3, pp. 289-309, [7]
<http://dx.doi.org/10.1080/19439342.2010.505027>.
- Lázaro, B. (2015), *Comparative study on the institutionalisation of evaluation in Europe and Latin America*, Eurosociál Programme, Madrid, http://sia.eurosociál-ii.eu/files/docs/1456851768-E_15_ENfin.pdf (accessed on 9 July 2019). [5]
- Ledermann, S. (2012), "Exploring the Necessary Conditions for Evaluation Use in Program Change", *American Journal of Evaluation*, Vol. 33/2, pp. 159-178, [9]
<http://dx.doi.org/10.1177/1098214011411573>.
- OECD (2020), *Institutionalisation, Quality and Use of Policy Evaluation, Lessons from countries' experience*. [2]
- OECD (2019), *Budgeting and Public Expenditures in OECD*. [16]
- OECD (2019), "Building a monitoring and evaluation framework for open government", in *Open Government in Biscay*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/a70e8be3-en>. [6]
- OECD (2018), *Draft Policy Framework on Sound Public Governance*. [1]
- OECD (ed.) (2017), *Systems Approaches to Public Sector Challenges*. [4]
- OECD (2016), *Open Government: The Global Context and the Way Forward*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264268104-en>. [10]
- Parkhurst, J. (2017), *The politics of evidence : from evidence-based policy to the good governance of evidence*. [11]
- Picciotto, R. (2013), "Evaluation Independence in Organizations", *Journal of MultiDisciplinary Evaluation*, Vol. 9/20, p. 15. [8]
- Rutter, J. (2012), *Evidence and Evaluation in Policy making*, Institute for Government, https://www.instituteforgovernment.org.uk/sites/default/files/publications/evidence%20and%20evaluation%20in%20template_final_0.pdf. [3]