

Improving decision making through policy evaluation in Belgium



Funded by
the European Union



OECD Public Governance Policy Papers
No. 31

Improving decision making through policy evaluation in Belgium

OECD Public Governance Policy Papers

This paper is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and the arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

This document, as well as any data and 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.

Authorised for publication by Elsa Pilichowski, Director, Public Governance Directorate.

Photo credits: Cover design by Meral Gedik using images from © thenounproject.com.

© OECD 2023

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.

Abstract

This paper provides an analysis of the Belgian federal government's practices regarding the institutionalisation, quality and impact of policy evaluations. The paper takes a wholistic approach, which not only looks at individual practices, but also at how those can come together so that evaluation becomes an integral part of the policy cycle.

Public policy evaluation promotes evidence-informed policymaking and that policies improve outcomes by bringing an understanding of what works, why, for whom, and under what circumstances. Evaluation is also a core tool of sound public financial management, as it helps governments spend better and promotes accountability and transparency in spending.

The paper proposes concrete policy recommendations for improving the Belgian federal government's evaluation system.

Acknowledgements

The paper was written by Claire Salama, Lead, Monitoring and policy evaluation, Public Management and Budgeting Division. The draft benefitted from inputs from Axel Mathot, Anne-Pauline De Clerc and Alessandro Ferrante.

This work was conducted under the leadership of Jón R. Blöndal, Head, Public Management and Budgeting Division, and the overall steering of Elsa Pilichowski, Director, Public Governance Directorate.

The action was funded by the European Union via the Technical Support Instrument, and implemented by the OECD, in cooperation with the Directorate-General for Structural Reform Support of the European Commission.

The paper was shared with the OECD Public Governance Committee and the Committee of Senior Budget Officials for comments.

Table of contents

Abstract	4
Acknowledgements	5
1 Introduction	8
2 Building a sound institutional framework for policy evaluation at the federal level	9
Introduction	9
A shared understanding of the objectives, tools and features of evaluation	10
Policy evaluation mandates	16
Cross-government evaluation frameworks	21
Summary of recommendations	23
3 Promoting the quality of evaluations: capacities, tools and practices	24
Introduction	24
Capacities for evaluation	24
Fostering skills for quality evaluations	28
Data for evaluation	32
Quality control and assessment mechanisms	37
Summary of recommendations	39
4 Ensuring that evaluations achieve impact	40
Introduction	40
Demand for evaluations	40
Embedding evaluations into key decision-making processes	43
Publicity and communication of results	45
Summary of recommendations	47
5 The role of BOSA in fostering a policy evaluation culture	48
Introduction	48
Co-ordinating evaluations across the federal government	48
A resources and structure	50
Summary of Recommendations	50
References	51

FIGURES

Figure 2.1. Governments' formal definition of policy evaluation	15
Figure 2.2. Institutions beyond the executive that have responsibilities related to evaluation	18
Figure 2.3. Institutions with competences related to policy evaluation across the government	20
Figure 2.4. Presence of a legal framework for policy evaluation across the government	21
Figure 3.1. Breakdown of the OECD OURData Index	35
Figure 3.2. Data value cycle in government	36
Figure 4.1. Different skillsets for using evidence	43

TABLES

Table 2.1. Federal ministries with a mandate to conduct policy evaluations	17
Table 3.1. Presence of a central unit in charge of evaluation in the ministry	26

BOXES

Box 2.1. Definition of policy evaluation	10
Box 2.2. Understanding the Belgian public governance system and its impact on the evaluation institution framework	11
Box 2.3. Spending reviews in Ireland	14
Box 2.4. National definitions of policy evaluation	16
Box 2.5. The Belgian Federal Planning Bureau	19
Box 2.6. The Netherlands Strategic Evaluation Agenda	22
Box 3.1. Skills for a high performing civil service	28
Box 3.2. The Belgian Civil Service Skills Framework	30
Box 3.3. Steunpunt Bestuurlijke Vernieuwing	31
Box 3.4. Data sources for analysis and evaluation	33
Box 3.5. The OECD OURData Index	34
Box 3.6. Access to administrative data in Denmark	35
Box 3.7. France Strategie's Evaluation guidelines	38
Box 4.1. The OECD evaluation criteria	41
Box 4.2. Skills for use of evidence	42
Box 4.3. Centralised evaluations portal in Norway	46
Box 5.1. Range of responsibilities of policy evaluation co-ordinating institutions	49

1 Introduction

Mobilising evidence for decision-making is crucial to achieving societal and economic goals, maintaining governments' resilience and preserving trust in public institutions. As countries face increasingly complex challenges – from climate change, to declining fiscal space and the multiplication of crises – evidence on what works is needed to guide government action. As such, evidence-informed decision-making contributes to preserving governments' agility and resilience in uncertain contexts. The covid crisis has also shone a light on structural and social problems, including the erosion of public trust in government and expert opinion. In a context of growing mis- and dis-information, governments' capacities to make evidence-informed decisions is thus vital to building citizens' trust in public institutions (OECD, 2022^[1])

Public policy evaluations can contribute to every government decision, from the design and implementation of long-term strategies, to budgets and resource allocation, and the impact assessment of new laws and regulations. (OECD, 2022^[2]). They also ensure that policies improve outcomes by bringing an understanding of what works, why, for whom, and under what circumstances (OECD, 2020^[3])

In the Belgian federal government, the importance of policy evaluation and its strategic use throughout the policy cycle is fully recognised in the current government agreement (Federal government of Belgium, 2020^[4]), and demonstrated in practice by the creation of the Directorate General for Budget and Policy Evaluation in the Federal Public Service for Strategy and Support. To harness the potential of policy evaluation for improving decision-making, the Belgian federal government can rely on good evaluation capacities in several line ministries and well-respected suppliers of evaluations.

However, overall, the Belgian federal government still faces many challenges that hinder an evidence-informed policy-making approach (Belgian Court of Audit, 2018^[5]). In particular, there is little co-ordination of evaluations across the federal government, and, as a result, limited opportunity for actors to share good practices, to conduct cross-sectorial evaluations in order to better understand the linkages and trade-offs between policy areas, to increase the overall production of evidence, and to create a critical mass of evaluation skills in the administration. Evaluation demand, both at a political level and at the level of the senior civil service, is also structurally low, thus affecting the impact of evaluative evidence on decision-making. This low demand ultimately impedes the incorporation of evidence into decision and policymaking and decreases incentives to strengthen evaluation mechanisms more broadly.

In this context, this paper provides an analysis of the Belgian federal government's evaluation system. It provides a gap analysis of its practices on the institutionalisation, quality and impact of policy evaluations by comparing them to OECD members' good practices in this regard. The paper thus takes a wholistic approach to policy evaluation, which not only looks at individual evaluative practices, but also at how those can come together so that evaluation becomes an integral part of the policy cycle. Based on this comparative analysis, the paper also proposes concrete policy recommendations for improving the Belgian federal government's evaluation system.

2 Building a sound institutional framework for policy evaluation at the federal level

Introduction

Embedding policy evaluation throughout the policy cycle helps integrate stand-alone or *ad hoc* evaluation efforts into more formal, frequent, and systematic approaches. Achieving this, however, requires creating a cross-government, institutional framework that provides incentives for conducting high-quality evaluations and can foster consistency – and thus comparability – of results across time and institutions (OECD, 2020^[3]).

There is no uniform approach to establishing a government-wide institutional framework for policy evaluation. The particular set-up for each country will depend on its national political and institutional context, as well as the objectives that it seeks to pursue through evaluation in the first place. Key governance choices can include, for example, whether to establish the framework in legislation and/or guidelines, or whether to allocate most responsibilities to a single co-ordinating institution or whether to grant more discretion to individual ministries.

Yet, empirical evidence from OECD countries shows that good practices can be identified (OECD, 2020^[3]). A sound institutional framework for evaluation includes:

1. A clear and shared understanding of the objectives, tools and features of evaluation.
2. Institutions (ministries or agencies, as well as actors outside of the executive) with the mandate and resources to conduct evaluations in their given field of expertise.
3. Institutions or actors with an explicit mandate to co-ordinate evaluations across institutions and to promote their quality and use across government
4. Cross-government high-level guidance, whether in a legal or policy-framework or in an evaluation agenda, for when and how to conduct evaluation.

This chapter discusses the extent to which the Belgian federal government integrates these different key elements of a sound institutional framework for policy evaluation.

A shared understanding of the objectives, tools and features of evaluation

Policy evaluation can be defined as the structured and evidence-based assessment of the design, implementation or impacts of a planned, ongoing or completed public intervention (see Box 2.1).

Box 2.1. Definition of policy evaluation

Policy evaluations are the structured and evidence-based assessment of the design, implementation or impacts of a planned, ongoing or completed public intervention, its design, implementation or results.

The concept of evaluation is multi-semantic. In some countries, evaluation can refer to the study of organisations, functions, procedures, policies, programmes or projects. The term evaluation is therefore sometimes used to refer to a continuum of tools aimed at improving the performance of policies, which include – but are not limited to – monitoring data and performance audits. These tools do indeed present common characteristics for evaluation. However, evaluation has a wider goal as it determines the relevance and fulfilment of a policy's objectives, coherence, efficiency, effectiveness, impact and/or sustainability.

The expression 'policy evaluation' is therefore preferred in order to focus on the evaluation of public interventions – i.e., the different means through which governments structure their efforts to attain desired objectives. Public interventions can be policies, programmes or plans.

In addition, this definition underlines that policy evaluation can refer to *ex post* evaluations, as well as *ex ante* and concurrent (or *ex durante*) analysis carried out when implementing policies to search for or improve means to reach the set policy objectives.

Policy evaluation can serve several complementary goals (OECD, 2020^[3]):

- *Policy planning*: evaluation, especially *ex ante*, allows decision-makers to assess an intervention's costs, benefits and potential impacts prior to its adoption.
- *Accountability and transparency*: evaluation can provide accountability on the use of public funds and resources by giving citizens and other stakeholders information on whether the efforts carried out by the government, including the allocation of financial resources, are producing the expected results.
- *Policy learning and improvement*: evaluation facilitates learning, as it clarifies why, how and for whom a policy was successful or not.

Similarly, evaluations can feed into different processes, such as budget decision-making, policymaking or strategic decision-making. Each of these goals and uses of policy evaluations are best served by a different institutional set-up and require different trade-offs. Therefore, to determine the best approach to setting up an evaluation system, one must first determine what objectives it pursues.

There is renewed interest in, and political commitment to, policy evaluation at the federal level in Belgium

Political interest in, and commitment to, policy evaluation is a strong catalyst for a sound government-wide evaluation framework. Indeed, mandates without the allocation of resources to conduct evaluations will remain ineffective, just as evaluation champions need some level of political support in order to provide incentives for ministries and agencies to follow their guidance.

The Belgian federal government has demonstrated a commitment to setting up the foundations of policy evaluations and more generally, to an evidence-informed approach to decision-making, through various initiatives in the past few years. First, the government agreement of 30 September 2020 references the need to conduct and/or use evaluations repeatedly (Federal government of Belgium, 2020^[4]). The government agreement is a key structuring document in the Belgian public governance system. This agreement lays out the agenda of each coalition government and is the result of a carefully negotiated consensus between the ruling parties. The coalition government agreement is thus the key document setting the federal government's policy agenda, which goes to show that policy evaluation is part of the current government's priorities (see Box 2.2).

Box 2.2. Understanding the Belgian public governance system and its impact on the evaluation institution framework

The Belgian public governance system is characterised by the following features:

- **Autonomous governments.** Belgian federal and federated (regional and community levels) governments have complete autonomy within their area of competence. The exclusive character of competences allocated to each level of government creates constraints on the extent to which they can benefit from the progress made in the area of policy evaluation at another level, in so far as laws or mechanisms adopted at the federal level will not extend automatically to the federated levels of government. This paper focuses on the Belgian federal government and its analysis and recommendations should be read with this in mind. The paper may refer to some good practices at the sub-federal level in some cases where they could be extended to the federal level as well.
- **Coalition governments and consensus-based decision-making.** The electoral system produces coalition government, and as a consequence, the political framework for policymaking is characterised by a search for consensus among coalition parties, acceptance of compromise and institutionalised power sharing. As a result, the government agreement, which is adopted at the beginning of each parliamentary mandate, serves as a strong compass for government action and reform.
- **Autonomous ministries within governments.** Within the federal government ministries are autonomous. This can generate challenges for the development and implementation of cross-government policy and rule-making tools and processes.
- **Pragmatism and informality in decision-making.** Consensus building within formal and often highly politicised structures, combined with the formal constraints imposed by the strict division of competences, can slow down and complicate the decision-making process. This is mitigated by a strong tradition of stakeholder dialogue.

As with any country, this local and political context affects the existing institutional set up for evaluation in Belgium. Addressing the different levers of a sound institutional system should take into account this context.

Second, the creation of the Directorate General Budget and Policy Evaluation within the Federal Public Service for Policy and Support (hereafter "SPF BOSA") through the Royal Decree of 22 February 2017, is further indication of the Belgian federal government's desire to strengthen the role of policy evaluation across government.

Some key federal decision-making processes rely on policy evaluation

The Belgian federal government has recognised the value of evaluations as decision-making tools by embedding their use in the regulatory and budgetary process. Since 2013, Regulatory Impact Assessment (RIA) is mandatory for all draft legislation, ministerial and royal decrees submitted to the Council of Ministers in Belgium (National Government of Belgium, 2013^[6]). RIA in Belgium consist of in an integrated evaluation of potential consequences and collateral impacts of regulatory projects on economic, social and environmental domains, structured around 21 themes covering the sustainable development goals, with a particular emphasis on gender, Small and Medium Enterprises, policy coherence for development, and administrative burdens.

The implementation, management, verification, and reporting of RIA are carried out by the Impact Assessment Committee (IAC). The IAC is an independent body composed of representatives from 5 administrations. Its secretariat is managed by the Agency for Administrative Simplification, responsible for the Better Regulation policies, and itself an independent service of the Chancellery of the Prime Minister.

Periodic *ex post* review of legislation is also mandatory for some legislation and sun-setting clauses are sometimes used at the federal level (OECD, 2021^[7]). *Ex post* assessments of regulations are systematic reviews of the stock of significant regulation against clearly defined policy goals, including consideration of costs and benefits, to ensure that regulations remain up to date, cost-justified, effective and consistent, and deliver the intended policy objectives. *Ex post* assessments complete the 'regulatory cycle' that begins with an *ex ante* assessment of proposals and proceeds to implementation and administration. Therefore, in theory, regulatory management tools such as *ex ante* RIA and *ex post* assessments of regulations can play an important role in supporting the use of evaluations and evidence in policymaking in Belgium.

Greater use of Spending Reviews is second mechanism by which the Belgian government has mainstreamed evaluation in decision-making, in this case in spending decisions. The Belgium federal government has begun implementing spending reviews on a more structured basis following the decision made by the Council of Ministers in December 2020. Spending reviews are tools for developing, assessing, recommending and adopting policy options by analysing the government's existing expenditure and linking this analysis to the budget process (OECD, 2022^[8]). The purposes of a spending review are to control the level of total expenditure, align expenditure to the priorities of the government and improve effectiveness within programmes and policies. Whilst traditionally spending reviews were used mainly to identify savings, they have been increasingly used to also improve policies more generally through better spending. In this regard, evaluations can and should feed into spending reviews, as they provide solid evidence on the efficiency, effectiveness and impact of expenditures.

By conducting evaluations of regulations through RIA or *ex post* assessments, as well as by conducting spending reviews, the Belgium federal government has demonstrated interest in an evidence-informed policymaking (EIPM) culture. Nevertheless, as the following sections will demonstrate, there are important gaps in the implementation of these tools and policy evaluations, which significantly affect the capacity for policy and decision making to be informed by evidence at the federal level in Belgium.

There is no shared understanding of what policy evaluations are and what objectives they pursue

Although Belgium explicitly recognises the importance of evaluation, it has not fully harnessed its potential as a tool for policy learning. At the federal level in Belgium, a shared understanding of the objectives and tools for evaluation does not exist and there is a bias towards evaluation for accountability and planning purposes, to the detriment of policy learning and improvement. This is in part due to the fact that evaluation activities in the federal government are:

- Either situated outside the executive, in the Supreme Audit Institution (Court of Audit) and the Parliament, where evaluations are a means to provide accountability to citizens on the government actions.
- or conducted *ex ante*, as part of the policy development process, with very little role for *ex post* evaluations within the executive.

Therefore, the potential for evaluations to provide lessons on what worked and what has not, why, for whom and under what circumstances, is not harnessed. The situation is compounded by the fact that there is no clear and shared definition of policy evaluation in the Belgian government. As a result, many actors conflate evaluation with a pure control and accountability tool. The development of certain policy processes that use evaluations are part of the continuum performance, such as RIA, which is often seen as a ‘check-the-box’ exercise submitted just days before the first approval of the Council of Ministers, or *ex ante* expenditure controls conducted by the Inspectorate of Finance, have also contributed to this situation. Yet, policy evaluation is a distinct tool with broader goals than that of control (see Box 2.1).

Policy evaluation involves analytical activities which determine, for example, the value of public intervention in improving the economic and social conditions of beneficiaries. It is different from other policy instruments, such as **monitoring**. While the latter checks progress against planned targets (answering the question “*has the target been met and to what extent?*”), policy evaluation looks at *how, for whom and why* the objectives of a policy have been achieved. Monitoring is of course an important step for *ex post* evaluation, but these tools are different from each other. Similarly, policy evaluation is different from **performance management**, which looks at the performance of individuals and organisations. Performance management aims to align and assess individual, team and organisational performance through agreed indicators and criteria which are regularly discussed and reviewed. Policy evaluation, on the other hand, does not look at the performance of organisations.

Similarly, policy evaluation is different from **audits** (including internal and external financial and compliance audits), which serve primarily as a control function. External audits by Supreme Audit Institutions such as the Court of Accounts serve to assess the reliability, transparency and accuracy of public entities’ financial reporting (financial audits) and to assess the compliance of public bodies’ financial statements with accounting rules and regulations (compliance audits) (OECD, 2011^[9]). Internal audits are the “*means by which the managers of an entity receive an assurance from internal sources that the processes for which they are accountable are operating in a manner which will minimize the probability of the occurrence of error, inefficient and uneconomic practices, or fraud*”, according to the International Organization of Supreme Audit Institutions (INTOSAI) definition (INTOSAI, 2019^[10]). Policy evaluation, on the other hand, is not only about accountability, but also about understanding why the objectives have or have not been achieved, how these results might differ across populations and time, and what can be done in order to improve the public intervention in question.

Policy evaluations also differ from **spending reviews**. Spending reviews are a structured and joint process led by the Ministry of Finance to identify, with the help of the line ministries, how existing expenditure can be better allocated to achieve the government’s priorities, as well as saving options, with the aim of feeding into the budget cycle. The goal of policy evaluations, on the other hand, is not strictly linked to feeding into the budget cycle. Policy evaluations also can assess any public intervention (i.e., projects, plans, as well

as programmes), as well as look not only at efficiency and effectiveness, but also at the relevance and fulfilment of a policy's objectives, coherence, efficiency, effectiveness, impact and/or sustainability.

Increasingly, spending reviews also seek to improve policies through better spending. In this regard, policy evaluations can feed into spending reviews. Indeed, given the relative short period in which spending reviews have to be conducted, it is important that they build on existing evidence, such as evaluations, but also performance information, on the topic that is under review. Because of this link between evidence (from evaluations and performance data) and good spending reviews, evaluations can be co-ordinated by the Ministry of Finance in some countries. As such, the Ministry of Finance encourages ministries to conduct regular evaluations of their policies, so that good evidence can be available to support the structured spending review process (see Box 2.3 for an example of how this process is managed in Ireland).

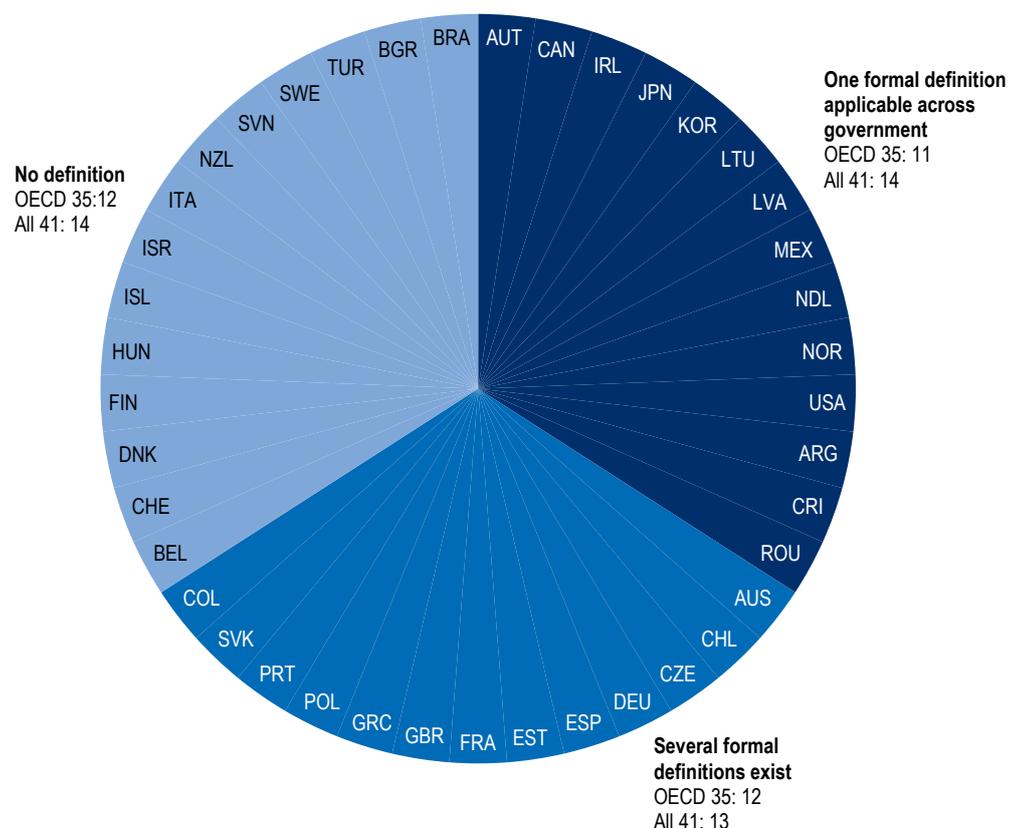
Box 2.3. Spending reviews in Ireland

In Ireland, the government's spending reviews aim to improve the allocation of public expenditure across all areas of government. The Department of Public Expenditure and Reform (DPER) is responsible for spending reviews and the way in which spending reviews integrate with the annual budget process. The Irish Government Economic and Evaluation Service (IGEES), established within DPER in 2012, supports the Department in this role.

Through the IGEES network, line ministries are systematically assessing the efficiency and effectiveness of government expenditures and policies through evaluations, both *ex post* and *ex ante*. DPER uses evidence from these evaluations, amongst other sources, to conduct a spending review process linked to the budget cycle, which considers in a more comprehensive manner the extent to which the government is achieving its goals through the expenditure committed on a given topic. The topics under review are identified for a two-year period (currently 2021-2022). As a result of this work, by October of each year, DPER publishes the spending review reports (one for each theme identified in the spending review period) in preparation for the fall budget discussions in Parliament.

Making the definition and the objectives of policy evaluation clear and communicating them would create a shared understanding for decision-makers, evaluators and citizens about the importance and purpose of this policy tool. This would also require making explicit what types of evaluations serve what purpose. To achieve this, a majority of OECD countries (23 out of 35) do have one (11 countries) or several (12 countries) definition(s) of evaluation (see Figure 2.1).

Figure 2.1. Governments' formal definition of policy evaluation



Note: n=41 (35 OECD member countries).

Source: OECD (2020^[3]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, <https://dx.doi.org/10.1787/89b1577d-en>.

Key conceptual terms in OECD definitions of policy evaluations typically fall across three categories:

- type of public interventions evaluated, such as policies, programmes, interventions
- criteria and goals for evaluation
- characteristics, such as quality attributes (systematic, objective, rigorous), timing (*ex post*, *ex ante*), and actors (internal vs. external evaluation).

These three categories are covered in the definitions of policy evaluations used in the Netherlands, the United States and Canada, for example (see Box 2.4).

Box 2.4. National definitions of policy evaluation

The Netherlands: “Policy evaluation is an examination of the efficiency (the extent to which the optimum effect is achieved with as few costs as possible and undesirable side effects) and effectiveness (the extent to which the policy objective is realised through the use of the policy instruments examined) of policy” (Ministry of Finance of the Netherlands, 2018^[11]).

United States: “Evaluation means an assessment using systematic data collection and analysis of one or more programmes, policies, and organisations intended to assess their effectiveness and efficiency.” (115th Congress, 2019^[12]).

Canada: “Evaluation is the systematic and neutral collection and analysis of evidence to judge merit, worth or value. Evaluation informs decision-making, improvements, innovation and accountability. Evaluations typically focus on programmes, policies and priorities and examine questions related to relevance, effectiveness and efficiency. Depending on user needs, however, evaluations can also examine other units, themes and issues including alternatives to existing interventions. Evaluations generally employ social science research methods” (Canada Treasury Board, 2016^[13]).

A first step, therefore, in clarifying the objectives and modalities of policy evaluation, would be for the Belgian federal government to adopt a definition of policy evaluation. This definition could also extend to the other levels of government, to ensure greater consistency in this practice in Belgium. Similarly, to other OECD countries, this definition could include what is policy evaluation, the type of knowledge it should produce, how and why it should be conducted and the actors that are involved.

In some OECD countries, the definition is embedded in a legal document. This is the case in Japan, for instance, where the Government policy evaluations Act (Act No. 86 of 2001) provides a definition of this policy tool. At the federal level in Belgium, however, in the absence of a legal framework for policy evaluation, such a definition could be included either:

- In guidelines on policy evaluations, that is, non-binding documents that provide guidance to ministries and agencies. The following section of this paper on the quality of evaluations offers additional information on what elements such guidelines could contain.
- Or in regulation, such as the SPF BOSA’s organisational decree. In Argentina, for instance, decree 292/2018 provides a definition of evaluation and designates the body responsible for preparing and executing the annual monitoring and evaluation plan for social policies.

Policy evaluation mandates

There is no single recipe for embedding policy evaluation across the government. Factors such as the political system, public administration cultures and the rationale for evaluation shape the development and characteristics of evaluation cultures and consequently how this is organised across government. Evidence from OECD countries shows that there is often a dual role in the evaluation institutional set-up: that of sectoral ministries and agencies, which carry out evaluations, and that of a (or several) central co-ordinating institution(s), which plays a key role in managing the evaluation eco-system by making sure that evaluation can take place at the right time and in the right place and that it can feed into decision-making (see Figure 2.3 for the type of institutions that can play this role).

Most ministries have the mandates to conduct policy evaluations

First, to ensure that policy evaluations are conducted, institutions, within and outside of the executive must have the mandate to conduct them. This is the case in the Belgium federal government where most ministries have a mandate to conduct policy analysis or evaluation. Indeed, policy evaluation is an integral part of most ministries' mandate (see Table 2.1). However, most ministries conduct evaluations on an *ad hoc* basis, in part due to a lack of a federal-wide framework for policy evaluations (see following sections).

Table 2.1. Federal ministries with a mandate to conduct policy evaluations

	Explicit mandate to conduct policy evaluation	Legal basis to the mandate
Chancellery (Administrative Simplification Agency and Impact Assessment Committee)	Yes	Law of 15 December 2013 and Royal Decree of 21 December 2013
Stratégie et Appui (BOSA)	No*	
Interior (Home Affairs)	Yes	Royal Decree of 26 April 2022
Foreign Affairs	Yes	Royal Decree of 25 April 2014
Finances	Yes	Royal Decree of 6 September 2018
Mobility and Transport	Yes	
Employment, Labor and Social concertation	Yes	
Public Health	Yes	
Economy	Yes	
Justice	No	
Ministry of Defence	No	

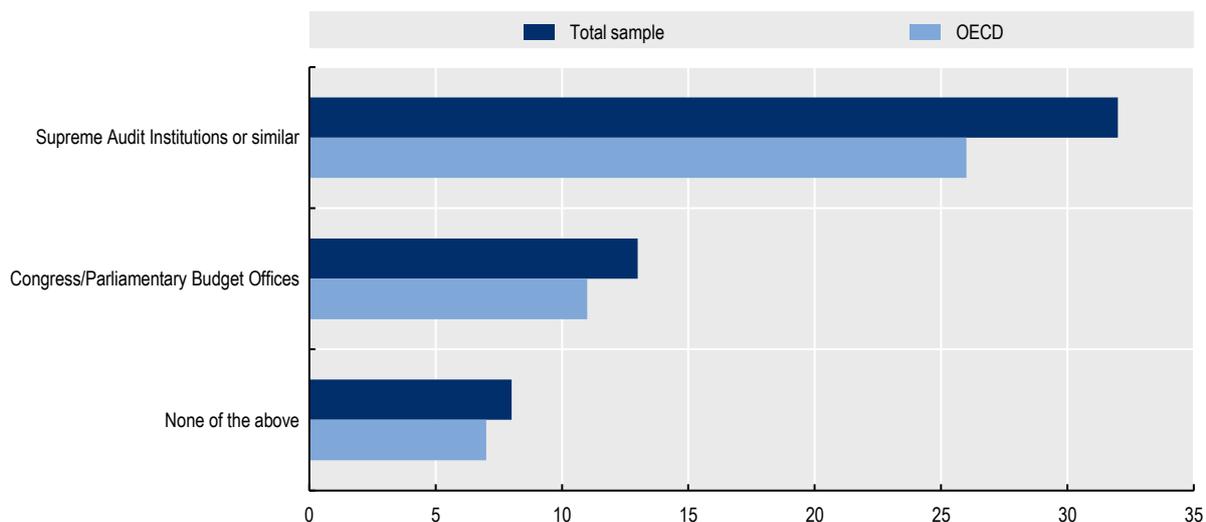
Note: The report to the King accompanying the Royal Decree of 22 February 2017 establishing the SPF BOSA clarifies that the policy support division within the DG for Budget and Policy Evaluation is responsible for "providing services that support the implementation of the integrated policy cycle". The DG Budget and Policy Evaluation is responsible for "offering services that will support the execution of the integral policy cycle even better. However, none of the 35 missions of the SPF BOSA (Article 2 of the decree) explicitly refers to support for policy evaluation.

Source: Authors based on OECD questionnaires and fact-finding mission.

The Ministry of Finance (known as "SPF Finance"), where the General Administration Expertise and Strategic Support unit is mandated to conduct studies and analysis on the outcome of policy options according to the royal decree of 19 July 2013 (National Government of Belgium, 2013_[14]). Similarly, the Ministry of Special Evaluation Office of the Foreign Affairs has a mandate to plan, execute and monitor evaluations related to development co-operation according to the royal decree of 25 February 2014 (National Government of Belgium, 2014_[15]). Moreover, each policy directorate within the Ministry of Health is responsible for policy support in his field of activity, including policy evaluation. Some ministries, however, do not have evaluation in their mandate. This is the case for the Ministry of Interior, for instance.

Several actors outside, or at arm's length, of the executive, conduct well-respected evaluations

There are also several actors outside or at arm's length of the executive with a firmly entrenched policy evaluation mandate. One of the missions of the Court of Audit since the law of 10 March 1998, for instance, is to conduct an *ex post* control of the economy, efficiency and effectiveness of federal expenditures (Belgian Court of Audit, 1998_[16]). As a result of this mandate, the Court conducts performance audits, much like is the case in most OECD countries (see Figure 2.2). These performance audits represent about one-third of the Court's work programme today.

Figure 2.2. Institutions beyond the executive that have responsibilities related to evaluation

Note: n=42 (35 OECD member countries). Answers reflect responses to the question, “Which of the following institutions beyond the executive have competences on policy evaluation at the central/federal level? (Check all that apply)”.

Source: OECD (2020^[3]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, <https://dx.doi.org/10.1787/89b1577d-en>.

Performance audits contribute to the accountability of government as they are sent to Parliament and require a formal response from the government. Their focus is decided independently by the Supreme Audit Institution itself (Belgian Court of Audit, 2021^[17]) which determines the scope of the audit, the audit questions, their methodology and timing. Although performance audits in general differ from policy evaluations, methodologies for both practices converge in many regards. In Belgium, this is especially true considering that the performance audits conducted by the Federal Court of Audit are increasingly focused on outcomes and impacts (Belgian Court of Audit, 2021^[17]). Indeed, while in the past this type of audit may have focused on the existence of preconditions for effective and efficient policy (design) and on goal attainment (process), recent examples demonstrate a shift toward more complex evaluations of outcome and impact of policy initiatives (Stockmann, Meyer and Taube, 2020^[18]). As a result of this practice, the Federal Court of Audit has become over the last few decades a well-respected actor in the evaluation field in Belgium.

The Parliament is also a source of policy evaluations at the federal level, although to a more limited extent. Indeed, it can ask for clauses to be included in legislation that include formal requirements for an *ex post* evaluation to be conducted of this law. However, at the federal level, the inclusion of evaluation clauses remains rather limited and follow-up to make sure that the evaluations are effectively conducted is even more rare (Stockmann, Meyer and Taube, 2020^[18]). The Federal Parliament can also request a (performance) audit from the Court of Audit.

Several institutions at arm’s length of the executive produce high-quality evaluations in the Belgian federal government. This is the case of certain research agencies, which are typically under the tutelage of a ministry but function with some degree of independence, with a separate budget line and autonomy over their work programme. One such research agency is the Health Care Knowledge Center (KCE), which is under the tutelage of the Ministry of Health. This research agency conducts research and analysis on topics related to healthcare, such as the organisation and financing of health care to the evaluation of medical technologies.

Another key actor in the policy evaluation playing field is the Federal Planning Bureau. The Federal Planning Bureau is an independent public agency that supports various Belgian public entities with

forecasts and policy-oriented studies on economic, social, and environmental issues. The Bureau is mandated to evaluate social and economic policies to enhance their rationality, efficiency, and transparency according to the law of 21 December 1994 (National Government of Belgium, 1994^[19]). The Federal Planning Bureau mainly provides *ex ante* studies, whether through macroeconomic forecasts or microsimulations, and a limited number of *ex post* impact evaluations (see Box 2.5). The Bureau is a very well-respected actor and is often solicited by government institutions for policy advice and analysis.

Box 2.5. The Belgian Federal Planning Bureau

Founded in 1959 as the “Programming Bureau”, the Federal Planning Bureau (FPB) is an independent public agency which, in its current form –as stipulated in the Law of 21/12/1994 amended by the Law of 04/04/2014– forecasts, studies and analyses policy measures. Its main mission is to support the political decision-making process. As such, it shares its expertise with the government, parliament, social partners and federal and inter-federal institutions. The FPB carries out studies and projections on economic, social and environmental policy issues, and on the integration of these policies within a context of sustainable development.

As a result of its mandate, the Federal Planning Bureau mostly focuses its work programme on *ex ante* evaluations and analysis, whether through macroeconomic forecasts or microsimulations. As part of its mandate, the FPB also provides the federal government, regions and communities with a common forecast (referred to as the “economic budget”) that serves as a starting point for their budgetary projections. The FPB also prepares a medium-term economic outlook including a detailed projection of general government accounts (federal, social security, federated entities and local) based on a no-policy-change scenario that is then used by the High Council of Finance (HCF) to make its recommendations.

The Bureau does conduct some *ex post* impact evaluations on an *ad hoc* basis, however, and has plans to increase this activity.

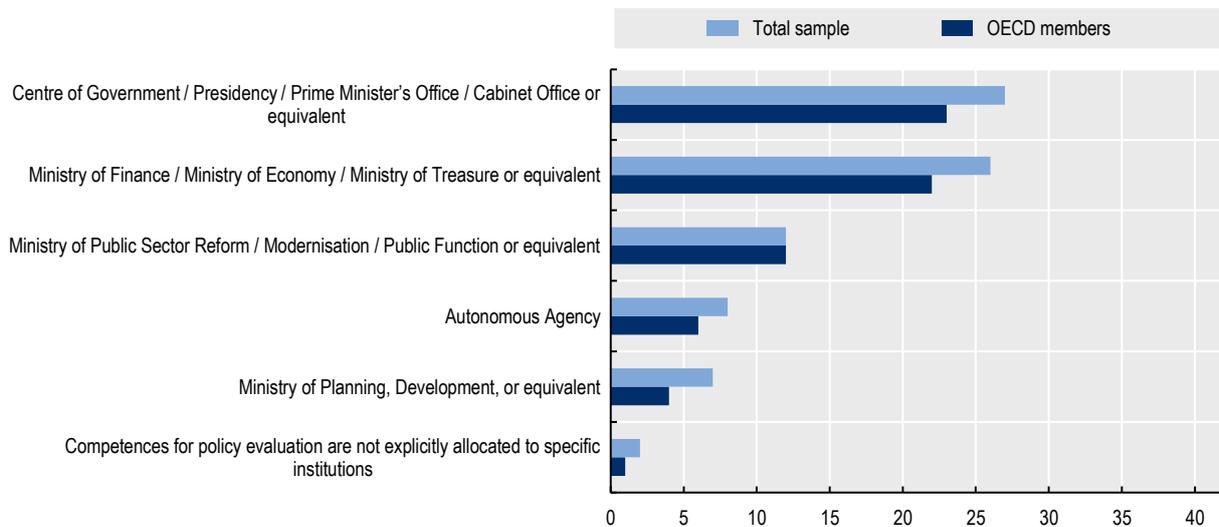
The FPB’s expertise is recognised throughout the government. This expertise is supported by its capacity to attract staff members with a high level of economic, econometric, and quantitative social sciences skills.

As a result of this expertise, the FPB is often solicited by other government institutions to analyse policy proposals and for advice, resulting in a heavy workload.

The lack of clearly identified evaluation champion(s)

Although the exact balance between centralisation and autonomy for line ministries in the evaluation system depends on the individual country, a degree of centralised co-ordination is a prerequisite for a strategic approach to evaluation. Indeed, a sound evaluation system entails the ability to look across a number of evaluations that cover different aspects of the same policy area to draw overall conclusions and identify common learning points.

This implies establishing evaluation champions, that is, institutions whose mandate it is to co-ordinate policy evaluations across the government. Evaluation champions are indeed needed to provide incentives to line ministries and agencies to conduct the evaluations, to co-ordinate evaluations across institutions and to promote their quality and use across government. OECD data shows that almost all OECD countries have at least one institution with this overseeing or co-ordination role (OECD, 2020^[3]).

Figure 2.3. Institutions with competences related to policy evaluation across the government

Note: n=42 (35 OECD member countries).

Source: OECD (2020^[3]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, <https://dx.doi.org/10.1787/89b1577d-en>.

In the Belgian federal government, there is no such clear co-ordinating institution for the moment. As a result, line ministries may be “searching for answers to the same issues” through evaluations, and at the very least, cross-sectoral evaluations that concern transversal policy issues are severely hampered (Stockmann, Meyer and Taube, 2020^[18]). The Federal Public Service for Strategy and Support (BOSA), has been given a role with respect to policy evaluation by virtue of its very name, but the scope of its mandate still needs to be defined more precisely as its organisational decree does not include any explicit reference to policy evaluation (Royal Decree of 22 February 2017 (National Government of Belgium, 2017^[20])). As a result of the absence of such a recognised and mandated champion, sectoral dynamics still prevail between line ministries and there is substantial variation across policy domains regarding the quality and frequency of evaluations.

Within the executive, mandates to conduct and use policy evaluations need to be further clarified

Beyond the role of these key institutions standing outside of the executive, the federal public services also need to have strong and clear evaluation mandates. Having a clear mandate, for example as expressed in an organisational decree, is the first step in ensuring that institutions dedicate resources and time to activity and that they do so on a systematic basis. All Federal Public Services could therefore have an explicit mandate to conduct policy evaluations, in their area of expertise, established in their organisational decree, as is the case currently for the SPF Foreign Affairs or the SPF Finance. These organisational decrees may also specify which unit or division is in charge of this activity (for more information on this, please see the following chapter on the capacities for evaluations).

In addition, there is much scope for greater co-ordination from SPF BOSA and for this administration to take on additional responsibilities related to the cross-government evaluation framework (see the last chapter on the role of SPF BOSA for more information on this topic).

Cross-government evaluation frameworks

While many institutions throughout the Belgian federal government do have evaluation mandates in their given field of expertise, and there is an understanding of the importance of evaluation at a high level, a clear and shared understanding of the objectives, tools and features of evaluation, high-level guidance on evaluation across government, and co-ordinating institutions are decidedly lacking.

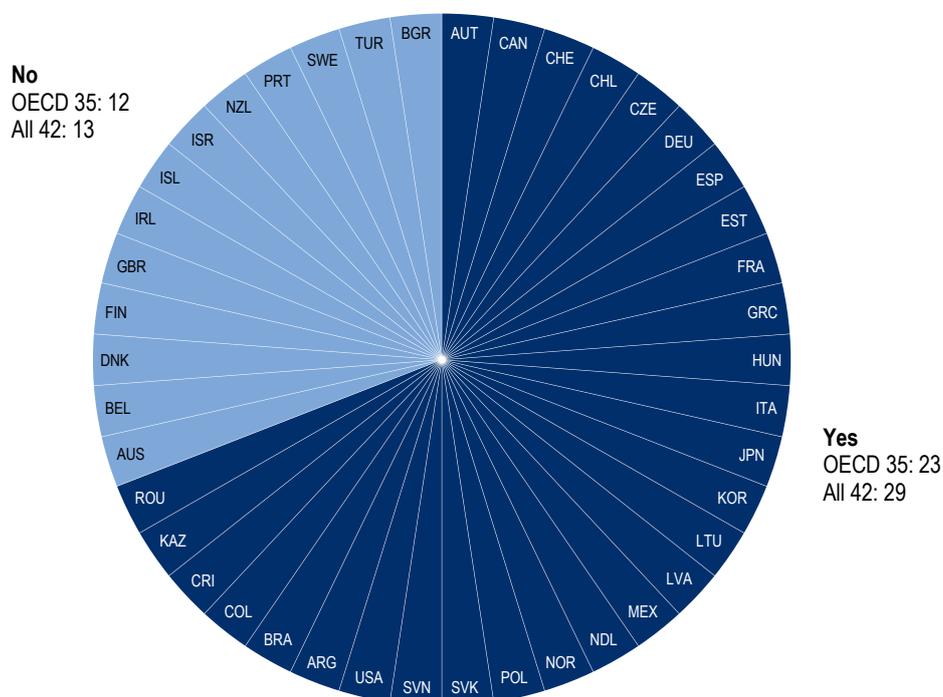
In particular, a government-wide framework providing guidance is crucial to systematise scattered evaluation activities, establish methodological standards, identify what type of evaluation is needed and when to feed into decision-making processes, and co-ordinate efforts for cross-sectorial evaluations.

There is no government-wide framework for policy evaluation at the federal level

At the federal level in Belgium, policy evaluation is not anchored in any type of government-wide framework. Clear guidance – for instance on mandates, timing and resources – would support the development of co-ordinated decisions and agreements on a common vision, mission and shared goals, which are all necessary steps in setting up a policy evaluation system.

An evaluation framework can be embedded into laws (whether the constitution, a primary law or secondary law), as is the case in the majority of OECD countries, or less formal policy documents such as evaluation agendas or guidelines. OECD data shows that a majority of countries (29 countries, 23 OECD countries) have adopted a legal policy framework to organise policy evaluation, either through constitutional provisions, primary legislation, or secondary legislation (OECD, 2022^[21]).

Figure 2.4. Presence of a legal framework for policy evaluation across the government



Note: n=42 (35 OECD countries).

Source: OECD (2022^[21]), "First lessons from government evaluations of COVID-19 responses: A synthesis", <https://doi.org/10.1787/483507d6-en>.

Empirical evidence suggests that, in those countries where a legal or regulatory framework for evaluation exists, it only defines the basic evaluation obligations of government departments and agencies and is complemented by more in-depth guidelines and planning requirements (OECD, 2020^[3]). In fact, evidence suggests that the existence of a legal framework in and of itself is not enough to sustain a sound evaluation system (OECD, 2020^[3]). A robust evaluation system needs to specify what should be evaluated, the actors that should be involved, their mandates, the timeline, the methodology and tools for evaluating, which is often done through non-binding policy evaluation guidelines.

Building a sound institutional framework for policy evaluations at the federal level in Belgium

At the federal level in Belgium, an institutional framework for policy evaluation is needed to:

- clarify the objectives (and definition) of policy evaluation,
- systematise evaluation activities across the administration by establishing clear mandates and providing institutions with resources for evaluations,
- identify what type of evaluation is needed and when to properly feed into decision-making processes,
- and to co-ordinate efforts for cross-sectorial evaluations.

To this end, the Belgian federal government could consider a three-tiered approach to setting up such a framework. In the short term, establishing such a framework would require making the SPF BOSA's mandate in regard to evaluation clear in its organisational decree to establish this actor as an evaluation champion across government, which promotes a culture of evaluation, co-ordinates cross-sectorial policy evaluations, and shares guidelines and good practices for how to conduct evaluations. See Chapter 5 on the SPF BOSA's mandate for more information on what this role could be. The framework should also clarify that the responsibility for conducting evaluations is with the line ministries.

In the medium term, the SPF BOSA could issue high-level guidance to ministries on why, how, and when to conduct policy evaluations in the form of a government-wide evaluation agenda, such as in the Netherlands (see Box 2.6), or a strategic plan. This document could be approved by the Council of Ministers.

Box 2.6. The Netherlands Strategic Evaluation Agenda

The Strategic Evaluation Agenda (SEA) aims to provide a structured approach to programming research and evaluation throughout the policy cycle. The SEA requires ministries to determine what type of evaluation evidence they will need and when over a 3-4 year period, to gather more relevant insights at the right moments for accountability and learning.

The SEA moreover also allows ministries to determine the best approach to obtain the necessary evidence for each policy area. Indeed, evidence can be gathered by means of a policy review, but also other types of *ex ante*, *ex durante* and *ex post* evaluations. In this way, the SEA is expected to lead to a better understanding during the whole policy cycle and the promotion of continuous improvement.

The programming of these activities largely follows the policy priorities from the policy agenda accompanying this budget. For the 2020 and 2021 studies and evaluations, concrete activities were launched in December 2020, for instance. The departmental SEA are sent to the Parliament annually and look forward 3-4 years, with the agenda for the next year being promised and for the following years an indication. The strategic evaluation agenda, however, offers the opportunity for line ministries to organise their evaluations on a broad thematic basis (not necessarily limited to a budget article) and to consider when they are most useful.

This guidance document should define a limited number of priority or cross-sectorial evaluations to be carried out in a given year across the government. The evaluations contained in the agenda or plan should be made public. Such an evaluation agenda does not preclude ministries from conducting policy evaluations out of their own initiative, of course, but rather ensures that key policy areas are covered and that resources are dedicated to cross-sectorial evaluations. In order to build the agenda, the SPF BOSA would need to consult with the line ministries and agencies in order to identify their most pressing needs in terms of evidence supply (i.e. what type of evidence on what works, why, how and for whom do they need for their policy development). These needs would be balanced against strategic policy priorities or for spending reviews, for example.

In the longer term, the Belgian federal government may consider adopting a legal framework to institutionalise policy evaluation beyond electoral mandates. Such a legal framework would focus on high level evaluation obligations of government ministries and agencies, as well as the need for an evaluation agenda or strategic plan. Details on when and how to conduct policy evaluations would be left to the accompanying agenda or strategic plan.

Summary of recommendations

- Adopt a comprehensive definition of evaluation applicable across the federal government. This definition could also extend to the other levels of government, to ensure greater consistency in this practice in Belgium. Similarly to other OECD countries, this definition could include what is policy evaluation, the type of knowledge it should produce, how and why it should be conducted and the actors that are involved.
- Establish evaluation agenda or plan, which specifies how many and which programmes and policies are going to be evaluated as a priority, who the evaluator is (whether they are internal or external to the administration), and when and how the evaluation should be conducted.
- Clarify roles, resources and responsibilities of the SPF BOSA in regard to policy evaluation. Update the Royal Decree of 22 February 2017 to include a provision in this regard (see chapter 4 for more details)
- Consider adopting one or several law(s) or regulation(s) framing the high-level evaluation obligations of government ministries and agencies in regard to policy evaluations. This law could include:
 - (i) A description of the different types of evaluation (e.g. *ex ante* or *ex post*; design, process and impact evaluations) to be carried out by the line ministries and is being evaluated (policy, programme, etc.).
 - ii) The resources (human and financial) dedicated to the evaluation.
 - (iii) The overall process and methods for the evaluation, including whom the results are sent to and how they are to be used (e.g. what decision-making processes their results should be used for).

3 Promoting the quality of evaluations: capacities, tools and practices

Introduction

Quality of evaluations is essential to ensuring their use for decision-making, and thus for evaluations to serve as tools for learning, accountability and better decision-making. However, quality is an important challenge faced by policy makers and practitioners in this area. This is due to a mix of capacity and skill gaps, issues with access and usability of data, heterogeneous oversights in the evaluation processes, and insufficient mechanisms for quality control (OECD, 2020^[3]). Yet, insofar as good quality evaluations benefit from greater credibility, they are likely to be given more weight in decision-making. The issue of the quality of evaluations is therefore a crucial one.

In order to promote the quality of their evaluations, the government can leverage three main assets:

- Public sector skills and capacities for evaluation,
- Data,
- Quality assessment (aimed at improving the evaluation methods and process) and control mechanisms (aimed at improving the evaluation product).

In this context, this chapter notes that many of the challenges linked to the heterogeneous supply of evaluation in the Belgian federal government can be traced back to issues with identifying and hiring staff with the appropriate skills to conduct policy analysis. This chapter also addresses the key role of data as a key enabler of quality evaluations and underlines the need for greater data accessibility and use in Belgium, as well as a more coherent approach to data governance for evidence-informed decision-making.

Capacities for evaluation

Human resources, in terms of capacity and competencies, as well as financial resources, have been identified as key enablers of policy evaluation by OECD countries (OECD, 2020^[3]). Indeed, when human resource capacities are limited, civil servants lack the time to conduct timely and methodologically rigorous evaluations, especially when those are not directly related to their operational priorities. Conducting policy evaluations, producing consistent data, and disseminating results can be financially and labour-intensive and require a critical mass of capacities in line ministries.

Capacities for evaluation exist in the federal government, with some policy areas being more invested than others

Capacities for evaluation do exist in the Belgian federal government. As mentioned previously, the Court of Audit is an important purveyor of evaluations in the Belgian federal context. Staff members split their work between compliance, financial and performance audits, as well as other tasks related to the Court's responsibilities. The Court estimates that about one-third of its 550 staff members work on performance audits.

Most ministries also do conduct evaluations somewhat regularly. Some good practices emerge in the administration in this regard. For instance, the development sector, due to its long tradition in the field of evaluation, benefits from important evaluation capacities. In addition to the evaluation capacities of the Belgian Development Agency (ENABEL), the Ministry of Foreign Affairs hosts the Special Evaluation Office, composed of 8 staff members and has an annual budget of 1 million euros. The Ministry of Finance also has a large team in charge of conducting policy analysis of fiscal measures, including supplying evaluations at the request of the policy divisions (the General Administration Expertise and Strategic Support has about 40 team members).

This observation should be nuanced, however, as most ministries outsource the large majority of evaluations that they conduct. This is the case even in ministries with higher evaluation capacities such as the Ministry of Finance. Outsourcing is a common practice in OECD countries, as evaluations require time, resources and highly technical skills that are not always commonly available inside the administration. Moreover, evaluations that are outsourced can benefit from an increased degree of independence and thus, perceived objectivity.

The challenge is that there is no general framework for outsourcing in the Belgium federal government, and outsourcing is managed in an *ad hoc* manner, regardless of whether the evaluation is attributed to universities or to private actors. Yet outsourcing policy evaluations to third parties requires specific commissioning skills (OECD, 2021^[22]). These skills are crucial to properly designing and oversee outsourcing contractual arrangements, and to manage and/or supervising evaluations (see the following section on skills). In the Belgium federal government there does not seem to be a clear understanding of what those skills are, nor guidelines established to set standards for the quality of external evaluations. As a result, ministries are not always in a position to assure the quality of the evaluations that they externalised, thus minimising their potential for impact on decision-making.

The heterogeneous organisation of the evaluation function in ministries means that capacities can be watered down and that the sharing of good practices is hindered

The organisation of the evaluation function is very heterogeneous across ministries in the federal administration. This is in part due to the varying needs of each ministry in this regard, but also to the absence of a common framework for evaluation. Indeed, in the absence of a clear and systematic mandate to conduct evaluations, not all ministries have found the need to dedicate specific and clearly identifiable resources to this task (see Table 3.1).

Table 3.1. Presence of a central unit in charge of evaluation in the ministry

	Central unit in charge of evaluation	Several units in charge of evaluation	No unit or dedicated staff for evaluation
Ministry of Finance	●	○	○
Ministry of Transport and Mobility	○	○	●
Ministry of Foreign Affairs, Foreign Trade and Development Cooperation ¹	●	○	○
Ministry of Interior	○	○	●
Ministry of Health	○	●	○
Ministry of Social Security		●	
TOTAL	2	2	2

1. The Special Evaluation Office of the Ministry of Foreign Affairs, Foreign Trade and Development Co-operation only evaluates development co-operation policies.

Source: Authors based on OECD questionnaires.

Only two ministries have reported having a central unit for evaluation and analysis. The Special Evaluation Office from the Ministry of Development co-operation is one of them. This office works as an independent entity in the Ministry of Foreign affairs, Foreign Trade, and Development Cooperation and is mandated to exclusively evaluate Belgian development co-operation policies and strategies, which are to a large extent developed and financed by the Directorate General for Development Cooperation (DGD). However, the Special Evaluation Office does not conduct the studies itself, but manages the outsourcing and supervising of the evaluations. The General Administration Expertise and Strategic Support in the Ministry of Finance, which conducts different studies on economic aspects of tax policy, is another one of these units. Finally, different policy directorates within the Ministry of Health host specialised evaluation units, but there is no central unit for the entire ministry.

Some ministries rely on a network of research agencies, which operate at arm's length of the ministry and therefore set their own agendas for analysis and evaluation. For example, an array of research agencies are under the tutelage of the Ministry of Health, including KCE, Sciensano, and the Superior Health Council. Although the Ministry of Health can influence its research programmes through its board members in these scientific institutes, the research agencies benefit from a large degree of autonomy in their research. The advantage of this setup is that their research and evaluations are seen as independent. The issue comes with the fact that the Ministry of Health cannot rely on its resources to meet its own analytical needs since the ministry does not have a direct impact on its work programme. Thus, whilst each directorate-general may call upon the research agencies to conduct some research on specific topics, they must also rely on internal capacities to conduct urgent *ex ante* analysis or more comprehensive reviews of their policies *ex post*. In the field of economic, social and environmental analysis, ministries can rest on the expertise of the Federal Planning Bureau similarly (see Box 2.5). For these topics, the Bureau acts alone in this role of research agencies for the central administration and thus is often overburdened, limiting its capacity to respond to all requests from ministries.

Generally speaking, though, the norm is for ministries to devolve evaluation activities to each policy unit or even to completely devolve evaluation activities to individual staff members. For instance, the Ministry of Health relies on policy analysis units present in each directorate-general such as the Data analysis unit in DG Healthcare. The Ministry of Interior relies on its staff members to conduct policy evaluations when it does not resort to outsourced contracts, but does not have any dedicated unit or team for evaluations. The Ministry of Transport and Mobility also relies on operational units that integrate evaluation activities into their daily work.

However, analysis and evaluation take time and expertise, which can be difficult to combine with having to handle daily and urgent tasks such as responding to parliamentary questions, responding to requests by the Government, or managing a project. As a result, staff in ministries might have little time or insufficient knowledge of evaluation techniques to conduct in-depth analytical work. This often results in a prioritisation of operational activities over evaluations, and the evaluation function is often performed by one or only a handful of civil servants who are strong believers in evaluation (Stockmann, Meyer and Taube, 2020^[18]). Furthermore, this means that evaluation practices can often be isolated from one another hindering the diffusion of good practices at a larger scale. Finally, when outsourcing practices are also spread across an administration with no well-defined guidelines and no co-ordinated approach, it prevents the building of structured evaluation networks.

Another important side effect is that few actors in the executive have a detailed understanding (or any understanding) of the resources that are dedicated to evaluation, and how these resources are funded. Policy evaluations are sometimes not even recognised as such, which leads to an underestimation of capacities at the federal level. This is compounded by the fact that, other than ENABEL and the Special Evaluation Office, no ministry has a separate budget line dedicated to evaluation activities. Yet, explicit budget provisions for policy evaluations are useful for several reasons. First, if funding for an evaluation is not determined at the inception of a policy, the analysis is more likely that it would be avoided or deferred (OECD, 2021^[22]). Moreover, not having clear and visible budget provisions for evaluations means that funding is likely to decrease as operational or urgent expenditures take precedent over the course of the budget cycle. As a result, several ministries reported in 2018 that financial resources for evaluation are decreasing in the Belgian federal administration (Belgian Court of Audit, 2018^[5]).

A critical mass of analytical skills is needed at the ministry level

Embedding proper evidence-informed decision making into government requires having a critical mass of analytical competencies available within ministries or Federal Public Services. Technical evaluation or analytical skills can be devolved to agencies or ‘research centres’, as is already the case in the Belgium federal government, and is commonly the case in Nordic countries, but governments also need a minimum level of core evaluation competencies within the administration itself.

Therefore, line ministries Belgian federal government should consider concentrating their evaluation capacities together as much as possible. For example, a range of countries, such as France, Canada, the United Kingdom, the Netherlands or Ireland have chosen to concentrate a significant mass of evaluation expertise within analytical units in Ministries. This model presents an important advantage as it improves how evaluation can be embedding into decision-making processes. Belgian ministries would undoubtedly benefit from having some critical mass of analytical skills in house, therefore.

In the shorter term, the Belgian federal government may consider mobilising research agencies attached to the ministries more systematically to provide *applied* research or analysis for evidence-informed decision-making processes, such as *ex post policy* evaluation but also regulatory impact assessment or value for money/effectiveness analysis for budgetary purposes). In parallel, an important step in ensuring that there is a critical mass of evaluation skills in ministries would be for the federal government to conduct an assessment or review of existing capacities and skills in this regard.

Fostering skills for quality evaluations

Conducting evaluations requires a certain set of skills and competencies, known more commonly as ‘policy analysis’ skills (see following sections on this topic). These competencies ensure or promote quality in evaluation practice, as individuals who possess the right competences are more likely to produce high-quality and utilisation-focused evaluations. In addition, public servants also need skills to use evaluations, in order to ensure that their results feed into decision-making. Skills for using evaluation and evidence are discussed in more depth in the following chapter.

There is a recognition that policy evaluation requires specific skills, which seem to be available in the Belgian job market

Simply put, a skill is “an ability to do something acquired through training and/or experience” (OECD, 2017^[23]). The OECD has developed a framework for civil service skills for public value, where analytical skills are one of four complementary and overlapping bundles of skillsets for a high performing civil service (Policy advisory skills). Box 3.1 provides further information on this framework and the different skillsets required for the civil service to deliver public value.

Box 3.1. Skills for a high performing civil service

The OECD report 2017 on civil service skills report identifies four main skill groups that are necessary to create public value:

- **Policy analysis:** These skills require leveraging technology and synthesising a growing range of evidence-informed scientific insights (e.g. behavioural economics, data science, strategic foresight) and a diversity of citizen perspectives for effective and timely policy advice to political decision makers.
- **Engagement:** these skills require working directly with citizens and users of government services to improve service experience, legitimacy and impact by leveraging the “wisdom of the crowd” to co-create better solutions that take into account service users’ needs and limitations.
- **Commissioning:** they involve designing and overseeing various contractual arrangements (outsourcing, PPPs, service level agreements, etc.) and managing projects to achieve impact through organisations (public, private, not-for-profit) that are best placed to deliver services due to their expertise and/or local position.
- **Network management:** they involve collaborating with a range of independent partners to address complex/wicked policy challenges by developing a shared understanding of the problem, collectively identifying potential solutions and co-implementation. While each civil servant does not need to be highly skilled in all of these areas, public institutions do require a solid mix of these skills in order to deliver public value in the modern public sector arrangement

Source: OECD (2017^[23]), *Skills for High Performing Civil Service*, <http://dx.doi.org/10.1787/9789264280724-en>.

Policy analysis skills or ‘analytical skills’, require that civil servants have the ability to generate and use robust and credible evidence. This includes an individual’s knowledge of different types of research methods, as well as fundamental skills of statistical and data literacy, and the capacity to read and understand analytical products. In practice, these skills often require a multidisciplinary set of competences drawing from a wide range of areas, including economics, statistics, social sciences, environmental sciences, law and engineering (OECD, 2021^[22]).

For these skills to be available to public administration, they must be either internalised by hiring staff members with those skills, or externalised, by tapping into these skills through knowledge networks, universities and the like. In some instances, evaluation skills can also be acquired through upskilling of existing staff members, although this proves more challenging given the complexity of the methods used.

There is a good supply of ‘evaluation skills’ in the Belgium job market, with several specialised courses on policy evaluation existing in federal universities. While there is no full-fledged master or Ph.D. degree in evaluation dedicated to policy analysis in Belgium, universities do propose degrees in quantitative economic and social sciences, as well as specific courses with a focus on policy evaluation. For instance, the Catholic University of Leuven (*Katholieke Universiteit Leuven*) is proposing a policy evaluation course of 6 ECTS that can be integrated into different academic curricula. Most other university-level initiatives are conducted by public administration departments, such as the evaluation trainings provided at the School for Public Administration for the Walloon and the French-speaking government (Stockmann, Meyer and Taube, 2020^[18]). There is also an inter-university certificate in public policy evaluation administered by the Catholic University of Louvain (UCLouvain), the Free University of Brussels (ULB), the University of Liège (ULiège), and the Walloon Institution for Evaluation, Foresight and Statistics (IWEPS).

The fact that evaluation requires specific skills is well-understood by the Belgian federal government. Administrations also do not hesitate to externalise evaluations to universities to conduct some more complex studies. The Belgian Research Action through Interdisciplinary Networks (BRAIN) programme managed by the Federal Public Planning Service Science Policy (Belspo), helped to develop the BELMOD project that aims to modernise the current microsimulation tool of the Ministry of Social Security Mimosis (Microsimulation Model for Belgian Social insurance Systems), for instance. This microsimulation tool analyses the impacts of social policy reforms, whether in terms of budgetary or distributional outcomes. During the COVID-19 pandemic, the University of Liège also worked on a study related to risk management with the Ministry of Interior and was commissioned to perform evaluations on cyber risks. Collaborations with universities take place on an *ad hoc* basis, however.

A limited number of institutions have also taken the initiative to conduct trainings for their staff members on policy evaluation methods. For instance, auditors of the Federal Court of Audit have access to in-house trainings with internal and external instructors on this topic. Junior auditors have access to a “beginners kit” with tools for public policy evaluation as part of their initial training package, while training for senior auditors may focus on case studies to further improve their evaluation skills. The Court of Audit also suggests to its auditors the possibility of taking a certificate in policy evaluation, such as the above-mentioned inter-university certificate in public policy evaluation. Staff members of the Special Evaluation Office also can follow trainings to strengthen the expertise of their collaborators in the evaluation field.

There is an overall lack of skills for evaluation inside ministries

However, overall, the Belgian federal government suffers from a lack of evaluation skills (Fobé, E., Peuter, B., Jean, M., & Pattyn, V., 2017^[24]). Analytical skills are key to ensuring the effective supply and use of evidence for decision making. In particular, quantitative skills, data skills and related soft skills are crucial in a world where the volume, velocity and variety of data has increased dramatically (OECD, 2017^[23]). Data scientists or economists/statisticians competent in working with data should be an integral part of ministerial staff to ensure that the evidence derived from data is used correctly, and that external evaluations and assessments are contracted appropriately (OECD, 2021^[22]).

First, the Belgian federal administration struggles to some extent with attracting and retaining ‘evaluator’ profiles. The salary, the length of the recruitment processes, as well as the lack of discretion when selecting candidates, attractiveness of the public sector. This may be dependent on the extent to which the administration is bound by the civil service framework for recruitment. The Ministry of Finance has been explicit regarding the difficulty it faces in recruiting and retaining staff members with high technical

evaluation skills. On the other hand, institutions such as the Court of Audit or the Federal Planning Bureau report that they do not face similar challenges.

Another important reason why Belgian federal administrations face issues with skills for evaluation is that there is no shared understanding of what these skills are and no mapping of related competencies across the federal government. Currently, the government relies on a basic competency model, with five groups of generic competences complemented by technical expertise (5+1 model, see Box 3.2). The competency model does not account for analytical or evidence-related skills.

Box 3.2. The Belgian Civil Service Skills Framework

The 5+1 model

The model is divided into 5 groups of generic skills and 1 complementary group comprising technical skills. The groups of generic skills are the following:

- information management
- task management
- collaborator management
- relationship management
- and personal development management.

Each group contains eight skills that are structured hierarchically based on their complexity and impact. Only the last group, personal development management, does not rely on this structure and present skills that are independent from one another. The technical skills group looks at civil servants' know-how, which is comprised of both expertise and the capacity to apply knowledge and skills in a given situation.

There are four categories to determine the level of proficiency of technical skill: basic, user, advanced, and specialist. There is not an exhaustive list of technical skills within the federal administration due to their large number and their specific nature.

For some very specific high-level senior positions, analytical skills or evaluation skills are an explicit requirement, but this remains a very rare occurrence. This was the case for example for the Director of General Budget and Policy Evaluation in SPF BOSA or the head of the Federal Internal Audit Service.

However, a general cartography of the existing analytical skills at the federal level is missing, preventing the administration from gaining a more granular understanding of skills and capacities in this area and thus establishing adapted recruitment strategies.

Apart from one-off training programmes proposed by some administrations, there is also no concerted effort to promote the professional development of the staff in charge of evaluation (vertical mobility) and the sharing of good practices between administrations (horizontal mobility). For instance, even in the Federal Planning Bureau, there is no concerted effort to provide policy and economic analysis training for the new recruits, notably due to time constraints. Therefore, applicants are expected to master all the necessary skills before joining this institution.

Co-operation with universities for evaluations happens on an ad hoc basis

When skills for evaluations are not present inside administrations, when certain evaluations require highly technical knowledge and competencies, or for reasons related to the independence of the evaluation,

outsourcing is an option. Governments typically outsource their evaluations to the private sector or to academia (universities), with some also involving Civil Society Organisations (although more rarely).

Apart from the BRAIN project, co-operation with universities for knowledge on evaluation happens on an *ad hoc* basis, with no structured agreement, whether at the individual level of each ministry or across the government. The *Steunpunt Bestuurlijke Vernieuwing* in Flanders (see Box 3.3) where, government and the universities agree on a research and evaluation programme for a five-year period, is a good practice for how to build more structured agreements between academia and government, in order to ensure that government can benefit from academia's expertise and research. An added value of these agreements is also that universities in Flanders work together in a networked setting, creating more synergies between their research.

Box 3.3. Steunpunt Bestuurlijke Vernieuwing

The *Steunpunt Bestuurlijke Vernieuwing* consists of a policy research agreement for a five-year period between the Flemish government and four different universities in Flanders: Ghent, Hasselt, Leuven, and Antwerp. Many benefits arise from this practice that aims to foster collaboration between universities and public actors. It provides stability for universities in their research programmes and also helps them to develop a network with the other universities involved. It also contributes to a culture of dialogue between the universities and the regional government which is more difficult to establish with short-term tenders. For instance, the design of the current research agreement 2021-2025 has been improved, with a stronger co-operation between both parties with respect to the previous version. It involves more co-creation exercises that allow for a better alignment of expectations and for a better equilibrium between scientific and policy-oriented components.

The Belgian government could consider adopting a government-wide approach to evaluation skills

While some ministries have taken to training their staff members in order to upskill existing personnel, the Belgian federal civil service has not developed a systematic government-wide approach in this regard and the training system of public servants is decentralised to individual ministries. This paper suggests a more systematic and government-wide approach to training staff on analytical (and particularly evaluation) skills. Nevertheless, the upskilling of existing staff in and of itself will not offer a structural solution to the existing analytical skill gaps in the Belgian federal administration.

To this end, the Belgian federal government could take the opportunity of the upcoming reform of the 5+1 Civil Service Framework to envisage creating a new analytical track within this framework, which could be managed centrally by the SPF BOSA in the longer term. This specific analytical track within the civil service could provide some flexibility in compensation, offer professionally attractive positions, with greater in-career horizontal and vertical mobility compared to other functions in the civil service. This analytical civil service track would be an integrated cross-government service, managed by the SPF BOSA, to support better evaluation across the civil service with analytical skills.

Several other OECD countries have created dedicated policy analysis tracks within the civil service, such as Ireland, and the United Kingdom or France. This system of analytical profession in the civil service contributes to making these skills available and visible in the public sector and ensures greater consistency of analysis and evaluations across the government, while facilitating mobility and exchange of good practices across ministries. Moreover, it provides a solution to the issue of public sector attractiveness. For instance, the IGEES has managed to develop name recognition in Ireland such that it is generally considered a more attractive career option than many other graduate programmes, including in the private sector, due to the horizontal and upwards mobility it provides early on in one's career (OECD, 2020^[3];

OECD, 2020^[25]). In addition to mobility, the attractiveness of the work is reinforced by the possibility of participating in quality seminars, in exchanging with peers, and in focusing on shaping high-priority policy initiatives. Moreover, the creation of such an analytical stream needs to be accompanied, as mentioned above, by a thorough exercise of analytical capacity mapping. Analytical resources and gaps have to be identified systematically to enable effective human resources management on this issue.

In the shorter term, the Belgian federal government should consider incorporating policy analysis skills within the revised civil service framework that SPF BOSA is currently working on. Indeed, the current 5+1 framework does not encompass these skills as they are defined in the OECD report on Skills for a high functioning civil service, nor in the new Joint Research Centre framework for competences for evidence-informed policymaking (OECD, 2017^[23]). This can create challenges in establishing a shared understanding of what these skills are, as well as in identifying existing skills in government.

Additionally, the Belgian federal government may wish to increase and systematise its co-operation with universities in order to encourage PhD candidates or researchers to work in the civil service for a few years in exchange for conducting some applied research, to harness the existing skills in the academic sector. These exchanges could take the form of a 'secondment' from the university to the Belgian federal administration. Other forms of co-operation could also look at how existing higher education programme in universities could feed into the future 'analytical' track of the civil service.

Data for evaluation

To produce reliable and robust analysis for evidence-informed policy advice, analysts in ministries need to have access to high-quality and timely data, as well as the appropriate tools and instruments to use this data.

Access to high-quality and timely data is needed to supply robust evidence

The quality and availability of data is a crucial challenge for evidence-informed decision making. In OECD countries, challenges related to access to data in the public sector generally include understanding what administrative data currently exist in ministries. There is also a broader data challenge that corresponds to the capacity of the public sector to generate the type of high-quality data that is necessary to produce evidence and evaluation (OECD, 2020^[3]). In other words, policy evaluation and evidence-informed policymaking (EIPM) can be hindered by:

- a lack of available data (see Box 3.4 for more information on what types of data are needed for evaluation)
- issues with data access
- and capacity gaps among government departments and agencies to generate data in a format that can be used.
- Data quality

Box 3.4. Data sources for analysis and evaluation

Conducting quality evaluation requires quality data, which may come from various sources:

- **Statistical data:** commonly used in research, it corresponds to census data or more generally to information on a given population collected through national or international surveys.
- **Administrative data:** this data is generally collected through administrative systems managed by government departments or ministries, and usually concerns whole sets of individuals, communities and businesses that are concerned by a particular policy. For instance, it includes housing data, tax records and data from public administrations.
- **Big data:** mainly drawn from a variety of sources such as citizen inputs and the private sector, big data is most often digital and continuously generated. It has the advantage of coming in greater volume and variety.
- **Evaluation data:** this data is collected for the purpose of the evaluation. It can take the form of qualitative questionnaires, on-site observations, focus groups, or experimental data. See further down for a description of impact evaluation methods to collect and analyse data.

Combining different data sources also has the potential to unlock relevant insights for policy evaluation. Applying big data analysis techniques to public procurement data can contribute to creating stronger, sounder and more relevant evaluations.

This understanding of the importance of access to data and the power of open data, exists in Belgium and some policy initiatives have been recently adopted, particularly in the field of open data. However, access to timely and quality data, particularly administrative data across ministries, as well as its use, remains an issue in the Belgian federal government today.

In the federal government in Belgium, different administrations have invested recently to increase the availability and accessibility of data

Availability and accessibility of data are important factors in data use, as data needs to exist but is also accessible to be used for analysis. Also, publicity of data matters as analysts may not otherwise be aware of existing data sets.

Recently, important efforts have been made in the field of open data at the federal level. The Belgian federal government has explicitly mentioned its intention to further open data across the federal administration in the last government agreement, as part of its intention to promote evidence-informed policymaking. The government strategy in terms of open data is led by the Directorate General for Digital Transformation in the SPF BOSA, in collaboration with other entities such as the Administrative Simplification Agency. Their approach is twofold: ensuring the availability of all administrative data at their disposal on a central repository website (data.gov.be), and actively convincing providers (other ministries) to adopt an open data approach. To ensure that they are capacitated to do so, the budget dedicated to the open data strategy has been increased for the year 2022. This directorate is also working with the ASA in implementing open data for better regulation, by facilitating access to different data sources relevant to the regulatory process.

Finally, since March 2021, the online platform PensionStat, created by Sigedis (public institution in charge of the management of the supplementary pensions data bank), the Federal Pensions Service, and the National Institute for the Social Security of the Self-employed, offers key figures on pensions in Belgium. Dashboards on different topics can be customised. Raw data can also be downloaded in Excel files.

In fact, data shows that Belgium has made continued progress since 2017 on pillar 1 of the OECD OURData index, which measures data availability (see Box 3.5).

Box 3.5. The OECD OURData Index

The OECD OURdata Index draws upon the OECD analytical framework for open government data policies. The OURdata Index is structured in three pillars which assess key elements of sound open data policies:

- **Pillar 1 Data availability:** *“measures the extent to which governments have adopted and implemented formal requirements to promote open government data at the central/federal level. This pillar also assesses how users are involved early on in the policy process as a means to inform data publication and identify policy needs. Pillar 1 also explores the scope of datasets available on central open data portals.”*
- **Pillar 2 Data accessibility:** *“measures the availability of formal requirements aiming at promoting unrestricted access to understandable data (e.g. open licence, metadata), the role of the ecosystem and the portal in ensuring data quality (e.g. feedback mechanisms), and the actual implementation of data accessibility requirements once these data are published.”*
- **Pillar 3 Government support for data reuse:** *“measures the extent to which governments play a proactive role in promoting the re-use of government data inside and outside government. This includes the definition and implementation of value co-creation initiatives and partnerships, capacity-building exercises, and governments’ efforts to monitor and evaluate policy impact.”*

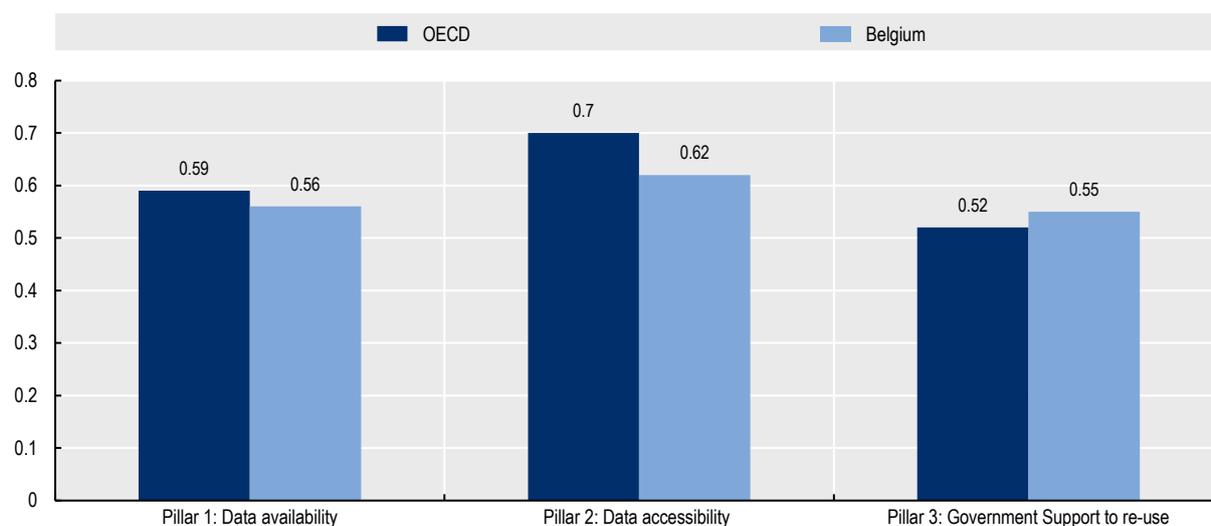
Source: OECD (2020^[26]), The OECD Open useful reusable data (OURdata) index.

There are also some good initiatives to promote data *accessibility* at the federal level in Belgium. Another project that will enhance the accessibility of data is the Health Data Authority project. This programme, conducted under the aegis of the Ministry of Health, aims to create a single point of contact for health care databanks. Within this project, the Ministry of Health is also collaborating with other health agencies on the creation of a Health Data Council, a co-operation platform between the different health public actors that would harmonise generation, management and exchange of data.

Another interesting example of the promotion of data accessibility and re-use in the Belgian federal government is the “data warehouse” of the Labour and Social Protection Market, which is managed by the Social Security Data Carrefour (BCSS). The BCSS links socio-economic data from various public social security institutions. Other than standard statistics that are calculated on a quarterly or yearly basis and that can be consulted via a web application, the warehouse also provides tailor-made data sets at the request of research and public authorities and supports projects such as the above-mentioned microsimulation tool BELMOD. Nonetheless, key challenges remain as access to the BCSS data can be characterised by long delays and the updating process of the data is not systematic.

A more proactive approach to data for evidence-informed decision making is needed

However, 2019 OECD data shows that Belgium is still lagging behind other OECD countries in regard to data accessibility and government support for re-usability. Indeed, Belgium was under the OECD average on the OURdata Index on accessibility and use of data (see Figure 3.1).

Figure 3.1. Breakdown of the OECD OURData Index

Source: Adapted from (OECD, 2020^[26]), OECD Open, Useful and Re-usable data (OURdata) Index: 2019, <http://www.oecd.org/gov/digitalgovernment/policy-paper-ourdata-index-2019.htm>.

First, there are some issues regarding accessibility of data, particularly when it comes to raw data. This is the case regarding the survey data collected by Statbel, as the institution only widely shares data that is published in its reports, i.e. secondary data. Such aggregate data cannot be merged or linked with other datasets, and thus are of little use for statistical analysis. Examples from other OECD countries, such as Denmark (see Box 3.6), suggest that greater availability of data does not have to be at the expense of its quality and of its potential for use. Whilst in theory, Statbel does make some anonymised data available for research and statistical analysis purposes (Statbel, 2022^[27]) stakeholders report that it is sometimes challenging for public institutions to access this data, and when they do, the data is anonymised in such a way that crossing of data sets is not possible.

Box 3.6. Access to administrative data in Denmark

In Denmark personal data is stored in registries with personal identification numbers. Statistics Denmark facilitates the use of these micro-level databases for research purposes for approved analysts, universities, research organisations or ministries. Statistics Denmark possess data in 250 subject areas ranging from labour markets, consumption, demographics to transport, agriculture and environment. The data is prepared by the Research Service Division and is accessible remotely and securely through specific internet servers. Analysts can access data in these areas as far back as from the 1970s.

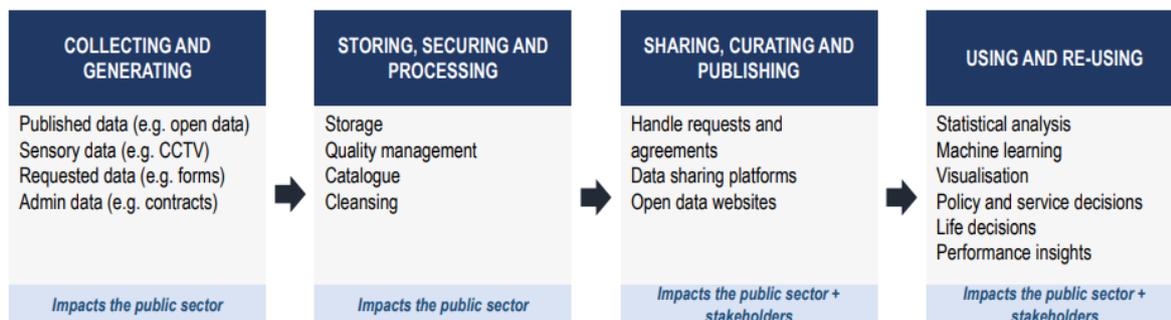
Moreover, the multiplicity of data integrators in Belgium might affect the capacity for public administrations to have an overview of available data sets and makes access to data tedious. At the federal level, there are several data integrators: DG Digital Transformation, Social Security Data Carrefour (BCSS), and e-health. Statbel also has a business register for statistical purposes, which, in addition to the data from the Crossroads Bank for Enterprises (which is managed by the FPS Economy), also contains other data such as data from the FPS Finance (VAT returns and corporate taxes), the National Social Security Office (NSSO), balance sheet data submitted to the National Bank of Belgium and data from the Quarterly Survey Data VAT Units, also from the National Bank of Belgium (NBB). In addition to these federal integrators, there are three regional integrators. This context might partly explain the absence of cartography of available datasets at the federal level. Furthermore, for the evaluation of some federal competences,

regional data is needed. Thus, evaluators wanting to analyse these policies would have to get authorisation for data access from each of the relevant integrators. Given that regions have a relatively important degree of autonomy, notably in areas related to open data, data for a given policy may not be open at all government levels, and when it is it may not be updated at the same frequency or be structured in the same way. This situation creates significant challenges for policy analysis, given that public governance in Belgium is characterised by significant interactions between the different levels of government on any given policy area. Having a single (or only a couple of) data registries across the Belgian government and at all levels of government may go a long way in promoting the supply of policy evaluation, as well as analysis more generally, in the country.

Some evaluators also report that the administration's interpretation of the European Union General Data Protection Regulation (GDPR) is too strict and hinders their ability to conduct analysis. Indeed, the General Data Protection Regulation, which was adopted by the European Union in 2016, aims to protect natural persons with regard to the processing of personal data and provides rules relating to the free movement of personal data. Given the importance of micro-data single identifier in the conduct of certain policy evaluations, a strict interpretation of GDPR might hamper data-sharing across the public administration. Due to the relatively recent implementation of this European legal framework, interpretation at the federal level has not been harmonised yet, which may lead to overly cautious approaches in some parts of the administration. Public research actors fear that this situation could result in restrained access to data for analysis compared to the period prior to the implementation of GDPR.

More generally, there does not seem to be a strategic approach to data production for analysis. An understanding of the data value cycle is needed (particularly from leadership) in order to embed a data and evidence-driven culture within the public sector, and ensure that public servants collecting and supplying data can think “use first” (OECD, 2021^[28]). The figure below provides a schematic illustration of this data value cycle.

Figure 3.2. Data value cycle in government



Source: Charlotte van Ooijen (2019^[29]), “A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance”; OECD (2019^[30]), *The Path to Becoming a Data-Driven Public Sector*, <https://doi.org/10.1787/059814a7-en>.

Thus, systematic strategies and policies to combine, link and re-use data, as well as to connect actors and decisions within and outside the public sector, are necessary to enable administrative data to be used for evidence-informed decision making (OECD, 2019^[30]). To encourage use and re-use of data, some OECD countries have developed explicit strategies to promote the use of data for evidence. The United States, for example, have institutionalised and implemented government-wide approaches to promote the use of data for analysis. They have done this by mobilising institutional resources, promoting internal champions and exploring the possibility of fully using existing data on a systematic basis through significant governance changes. The United States have issued the 10-year Federal Data Strategy centred around 3 core principles (ethical governance, conscious design and a learning culture), which is accompanied by

the implementation plan of 40 practices that help agencies to comply with the Federal Data Strategy (OECD, 2019^[30]). This Data Strategy is complemented by the Foundations for Evidence-Based Policymaking Act of 2018, which includes a government-wide approach to data as a key pillar for evidence-informed policymaking. Such a strategic approach is lacking in the Belgian federal government.

The Belgian federal government could thus adopt a common data governance strategy to better support data for evaluation and evidence-informed decision making. This could include a systematic mapping of registries, administrative data and surveys, a central portal for making data available for public use, and a specific process for facilitating access to merged anonymised files, under specific authorisations for government institutions, universities and researchers.

Quality control and assessment mechanisms

To be robust and trustworthy, evaluations need to be independent and methodologically sound. To this end, governments can put in place mechanisms to ensure good governance and the methodological robustness of policy evaluations. These can be either quality assurance mechanisms, which seek to ensure credibility in how the evaluation is conducted (the process), or quality control mechanisms which instead focus on the end product (the report) ensuring that it meets a certain standard for quality (OECD, 2020^[3]). A mix of quality assurance and quality control mechanisms is very important to really foster the quality of evaluations.

In the Belgian federal government, there is no formal process to ensure the methodological quality of policy evaluations in general. Regulatory Impact Assessments can be submitted to the Impact Assessment Committee (IAC), which provides advice on RIAs upon request by the responsible ministry and reports annually on the quality of RIAs and functioning of the RIA process. But the IAC is only involved at the request of the responsible ministry and, in many cases, time constraints mean that this is not the case.

Other than the IAC, government-wide mechanisms to improve the quality of evaluations, such as peer reviews, evaluation supervisory groups, evaluation guidelines and standards, or checklists, do not exist. For instance, evaluators report that they refer to international evaluation guidelines when needed. Evaluation guidelines are nonbinding documents or recommendations that aim to support governments in the design and implementation of a policy and/or practice (examples include white books and handbooks). They can address a wide variety of specific topics including the design of evaluation approaches, the course of action for commissioning evaluations, planning out evaluations, designing data collection methods, evaluation methodologies or the ethical conduct of evaluators. Whilst international guidelines may prove useful for individual evaluators to refer to, there is a risk that not having federal guidelines in Belgium can reinforce the heterogeneity of practices across the administration. Furthermore, guidelines developed by international organisations, such as the World Bank Group's Evaluation Principles, are mainly aimed at ensuring that the evaluations that their subsidiary institutions conduct are technically robust and credible (World Bank Group, 2019^[31]). As mentioned in chapter two of this paper, an evaluation's objectives influence the way that it should be set up, managed and conducted. There is therefore a limit to the transferability of such guidelines to the federal context where evaluations must feed into specific decision-making processes.

Similarly, some administrations do seek to conduct peer reviews their evaluations by collaborating with academics or independent experts in this regard. These peer reviews, however, occur on an *ad hoc* basis and are largely based on personal relationships between the evaluator/ evaluating team and the reviewers.

The SPF BOSA could adopt federal policy evaluation guidelines

An important step in promoting the quality of policy evaluations across the federal government will be to adopt methodological guidelines on how to conduct, commission and use evaluations for all government institutions.

Guidelines can provide general guidance on what types of evaluations are needed, as well as how to conduct evaluations from a theoretical and methodological standpoint but should also be tailored to public servants and thus contain very practical information in this regard. For example, guidelines can specify what steps to take to commission evaluations, from how to define the terms of references to how to issue a call for proposals. A good practice in this regard is that of the United Kingdom, which has developed several cross governmental guidelines, the most important being the Magenta and Green Books. France Stratégie's evaluation guidelines also provide an interesting example, which presents different methods for conducting robust evaluations (see Box 3.7). As an evaluation champion, the SPF BOSA will be best placed to develop and issue similar guidelines.

Box 3.7. France Strategie's Evaluation guidelines

France Stratégie is a French agency attached to the Prime Minister but operating at arms' length from the government that provides expertise on major social and economic issues through *ex post* evaluations of public policies, analysis notes, debates and consultative exercises. The institution has issued guidelines on impact evaluations to help decision makers and practitioners conduct evaluation and analyse evaluation results (Comment évaluer l'impact des politiques publiques: un guide à l'usage des décideurs et des praticiens (France Stratégie, 2016^[32])).

Firstly, these guidelines present different methods for conducting scientifically reliable impact evaluations to deduce a causal relationship between the public intervention being evaluated and the effects it has on its beneficiaries (relevant indicators include health, employment, education, etc.). Several methods are explained in detail (including 'differences in differences' and randomised controlled trials), and the guidelines emphasise - among others - the importance of building a credible counterfactual, of choosing relevant indicators, and of avoiding selection mechanisms that could skew the results of the evaluation. In the next section, the guidelines address the question of how to analyse evaluation results and identify the reasons for the success or failure of a public policy. Finally, the guidelines explain how to compare the effects of multiple policies with the same goal and choose the most efficient one. This last section covers cost-benefit and cost-effectiveness analyses.

France Stratégie's guidelines are concrete and user-oriented, as they take into account a wide range of scientific and operational constraints that often surround the implementation of evaluations. Such constraints include the availability, breadth and quality of data and the evaluation budget. These constraints often determine the evaluation method to use, as different methods require different types of data (for instance, the matching method requires rich data on individuals and their social and economic environment).

Finally, France Stratégie's guidelines are policy-oriented and help decision makers bridge the gap between policy evaluation and decision making, as they give clear recommendations on how to use evaluation results to improve public policies, and how to strengthen the evaluation capacities of policymakers. The guidelines recommend, inter alia, conducting systematic reviews of the existing evidence in order to assess whether evaluation results converge and diverge depending on the institutional context of the policy. The need for policy makers to institutionalise and operationalise the production and access to data is also emphasised in the guidelines.

Other than providing information on how to conduct high-quality evaluations, the guidelines should also specify what steps to take to ensure the ‘good governance’ of the evaluations, that is to say to ensure their independence, as well as the ethical conduct of the evaluators.

Summary of recommendations

- Create a critical mass of evaluation and analytical skills in line ministries:
 - Consider concentrating the evaluation capacities of ministries in one place (whether it be one unit or division, or within a larger team focused on a variety of analytical skills such as data, statistics, evaluation, regulatory impact assessment, etc.).
- Adopt a government-wide approach to evaluation skills:
 - Adopt a more systematic and government-wide approach to training staff on analytical (and particularly evaluation) skills by conducting more regular training, identifying skill gaps and identifying key factors to fill those gaps.
 - Increase and systematise, in the shorter term, co-operation with universities in order to encourage PhD candidates or researchers to work in the civil service for a few years in exchange for conducting some applied research, in order to harness the existing skills in the academic sector.
 - Use the upcoming reform of the 5+1 Civil Service Framework to consider creating a new analytical track within this framework, which could be managed centrally by the SPF BOSA. This specific analytical track within the civil service could provide some flexibility in compensation, offer professionally attractive positions, with greater in-career horizontal and vertical mobility compared to other functions in the civil service. This analytical civil service track would be an integrated cross-government service, managed by the SPF BOSA, to support better evaluation across the civil service with analytical skills.
- Adopt a data governance strategy to better support data for evaluation and evidence-informed decision making:
 - Develop clear guidelines for use of personal data for research purposes, as well as the limits to such use in relation to the application of the European General Data Protection Regulation.
 - Merge data registries, where possible, across the Belgian federal government and at all levels of government.
 - Support data production in line ministries for evaluation by identifying evidence needs and data production.
- Adopt policy evaluation methodological guidelines across the federal government on how to conduct, commission and use evaluations for all government institutions.

4 Ensuring that evaluations achieve impact

Introduction

Effective use of evidence and evaluations is key to embedding them in policy and decision-making processes. Without use of evaluations, gaps will remain between what is known to be effective and decision making in practice. Moreover, as policy makers invest public funds in conducting evaluations in the hopes to improve policies and programmes and provide useful insights on public issues, their use is key. On the other hand, underuse of evaluations can jeopardise the evidence-informed decision-making agenda. When decision makers ignore the results of evaluations, for instance, future calls for evaluation may be undermined, and evaluations become check-the-box exercises.

Unfortunately, conducting evaluations is not a sufficient condition for them to be used. Demand from primary intended users also needs to be there. Yet, demand for evaluations by decision-makers is often limited. Ensuring that evaluations achieve impact and are actually used is thus a complex task that involves a variety of overlapping factors. First, the impact of evaluations depends on environmental factors linked to the general public governance system, the institutional culture and external pressures for greater use of evidence in decision-making play an important role in this regard. Governments can also incorporate evaluations into key decision-making processes, such as budgetary or regulatory processes, to promote their systematic use. Finally, factors such as the publicity and communication of evaluations, as well as civil servants' skills for using evidence, come to play.

In the Belgian federal government, use of evaluations and analysis remains an important challenge. This chapter looks at the different ways in which the Belgian federal government could promote more systematic connections between supply of, and demand for, evaluations, either through increased publicity and communication of results, or through policy frameworks that embed use of evidence in key decision-making processes.

Demand for evaluations

Cultural and societal factors may affect the extent to which evaluations gets used in decision making (OECD, 2020^[3]). For instance, societal attitudes towards policymaking, and what and who should contribute to it, can also affect the use of evidence.

In a country such as Belgium that is highly fragmented along linguistic, regional and political lines, political parties play a decisive role in the policy-making process. Government agendas are set through government agreements, which are the result of careful consensus between the ruling parties, and there is little margin for ministers and administrations to formulate policies outside of this framework. In addition, whilst ministries may be asked to provide inputs for the text, the governmental agreement is principally a political

document resulting from political negotiations. As a result, there is little room for use of evaluations in policy formulation and in evidence-informed decision-making processes may suffer.

Therefore, ministries are often limited by the agenda set by the government agreement, and the public administration's scope in formulating policy proposals is limited. This means that demand for evaluation and evidence, outside of the key moment of negotiations over the government agreement, is inherently limited.

Still, there is some demand from policy makers for evaluations. This is especially the case in some policy areas where use of evidence is part of a long-standing culture. In the health sector, for example, the need to rely on scientific evidence for policy formulation is evident and the ministry relies on advice from Scienscano and other research agencies regularly.

For most policy sectors, however, demand focuses mostly evidence-informed policymaking is limited to the use of RIA, and – as of very recently – spending reviews. Whilst these types of evaluations play key roles in planning policies (*ex ante* impact assessments) and assessing the efficiency and effectiveness of expenditures (spending reviews), more systematic evaluations of policies' outcomes in terms of relevance, coherence and impact, are also needed. This means that administrations need to look at the full range of public interventions, and focus on the entire spectrum of evaluation criteria (see Box 4.1).

Box 4.1. The OECD evaluation criteria

The OECD DAC Network on Development Evaluation (EvalNet) has defined six evaluation criteria – relevance, coherence, effectiveness, efficiency, impact and sustainability – and two principles for their use. These criteria provide a normative framework used to determine the merit or worth of an intervention (policy, strategy, programme, project or activity). They serve as the basis upon which evaluative judgements are made.

The criteria can be defined as follows:

Relevance: *The extent to which the intervention objectives and design respond to beneficiaries, global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.*

Coherence: *The compatibility of the intervention with other interventions in a country, sector or institution.*

Efficiency: *The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.*

Effectiveness: *The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.*

Impact: *The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.*

Sustainability: *The extent to which the net benefits of the intervention continue, or are likely to continue.*

Source: OECD (2021^[33]), *Applying Evaluation Criteria Thoughtfully*, <https://doi.org/10.1787/543e84ed-en>

Moreover, in the Belgian federal administration especially, political requests for *ex ante* analysis and modelling tends to crowd out *ex post* evaluation studies that look at actual outcomes and impacts. For example, most of the studies produced by the General Administration Expertise and Strategic Support from Ministry of Finance are *ex ante* evaluations requested by the political level. This is also the case for the knowledge or research centres that exist in the federal government. As discussed previously, the work

of the Federal Planning Bureau is more balanced towards *ex ante* evaluation (as per its mandate). Indeed, the high level of expertise of the Bureau on an array of socio-economic questions means that it is often solicited to measure the potential impact of policies, leaving little time for studies that are conducted at the Bureau's own initiative.

Yet, not only are *ex post* evaluations just as crucial as *ex ante* assessments, it is important for evaluation units or teams to preserve a certain degree of autonomy in choosing what interventions they will evaluate and according to what evaluative questions. This promotes the credibility of evaluations, and ultimately their quality of studies (OECD, 2021^[22]).

An investment in skills and proper planning can promote demand for evaluation evidence

In order to ensure that this autonomy is possible, and that the evaluations that are conducted can properly feed into decision making in a real way (and not a 'tick-the-box' formalistic way), policy evaluations need to be planned early. This requires, on the one hand, having an evaluation agenda or plan when it comes to *ex post* assessments, and, on the other hand, proper legislative planning to anticipate analytical needs over the course of a government mandate.

Furthermore, stimulating demand for evidence requires behaviour changes from decision makers and senior leadership, which can be promoted by developing certain skills in this population. The OECD together with the European Joint Research Centre (JRC) has identified 6 clusters of skills that the public sector should aim at developing in order to employ the evidence throughout the policy cycle. Box 4.2 expands on this skillset and provides definitions for each cluster of skills based on this joint OECD-JRC work.

Box 4.2. Skills for use of evidence

This skill-set is defined as a collective skill-set for the improvement of public service in the future and not as a full list of skills that each public servant needs to master. This skillset does not apply to one scenario; instead, it is of a cross-cutting character and can be applied on multiple occasions. It includes elements like critical thinking, systems thinking, and engaging with stakeholders.

- **Understanding EIPM** – understanding of the policy cycle and knowing how evidence could be employed in each of its component. It has to be underpinned by the familiarity with the fundamental methods in research and statistics.
- **Obtaining Evidence** – ability to recognise and measure the existing stock of evidence in the relevant policy area and identify the evidence gaps to commission high quality studies.
- **Interrogating and Assessing Evidence** – ability to assess the provenance, reliability and appropriateness of evidence by using systemic, holistic and critical thinking tools free of personal bias.
- **Using and Applying Evidence in Policymaking** – deep knowledge of the policy area and understanding how different evidence, research and innovative approach can be used to support policy design and implementation.
- **Engaging with Stakeholders in EIPM** – strong engagement and communications skills. Ability to engage various groups of stakeholders in a discussion and to communicate policy messages effectively.
- **Evaluating the Success of EIPM** – ability to use different evaluation approaches to inform and improve EIPM processes and policy cycle.

Figure 4.1. Different skillsets for using evidence



Source: Adapted from OECD (2020^[34]), *Building Capacity for Evidence-Informed Policy-Making*, <https://doi.org/10.1787/86331250-en>.

OECD country practices aimed at improving these skills reveal a wide range and approaches towards skills development interventions. The work by the OECD on Building Capacity for Evidence Informed Policymaking (OECD, 2020^[34]) suggests that training for Senior Civil Service leadership should be aimed at increasing managers' understanding of evidence informed policymaking and policy evaluation, enabling them to become champions for evidence use. In Canada, for example, the executive training in research application (EXTRA) programme provides support and development for leaders in using research. Intensive skills training programmes aimed at policy makers may be more focused on interrogating and assessing evidence and on using and applying it in policymaking. The SPF BOSA could thus consider organising training for Senior Civil Service leadership on these issues. Such trainings can take the form of workshops, masterclasses or seminars. These skills are indispensable for the institutional and cultural transformation that is necessary to foster the demand for evidence at the senior civil service and political levels.

Embedding evaluations into key decision-making processes

Formal organisations, institutional mechanisms and processes set-up a foundation for evidence-informed decision making that can withstand transitions between leadership (Results for America, 2017^[35]). Indeed, formal decision-making processes create a fertile ground for supply and demand for evaluations to meet.

These mechanisms can be found either at the level of specific institutions, such as management response mechanisms (i.e. when management is required to provide a formal response to the evaluations it has requested), or within the wider decision-making cycle, such as through the incorporation of policy evaluation findings into the budget or regulatory cycle or discussions of evidence in strategic planning (OECD, 2020^[3]).

At the federal level in Belgium, some of the frameworks exist, or have recently been implemented, but others need to be improved in order to support more systematic impact for evaluations.

Systematic links between evaluations and decision making remain limited

Some types of evaluations can feed key decision-making processes. This is the case of the performance audits done by the Court of Audit, which are sent to Parliament. The regulatory process also makes use of evaluations in theory. Finally, the spending review process, which has recently been initiated at the federal level in Belgium, can make use of policy evaluations. Although this has not yet been the case in the first pilot spending review, it is to be expected that evaluations will feed into the coming rounds, since these reviews take into consideration existing analysis such as evaluations. Furthermore, the analysis resulting from spending reviews will also benefit future evaluations. This remains to be seen in practice, however, at the spending review process is still in its early stages at the federal level in Belgium.

Other than this, there are *ad hoc* examples of evaluation recommendations that are taken up by decision makers. For instance, the study conducted by the Federal Planning Bureau and the Court of Audit on the impact of reduced employer social security contributions on first recruitments induced the government to review the initial measure (Belgian Court of Audit, 2021^[36]). Another example concerns the study made by Court of Audit on the efficiency of public policy on social and tax incentive within the supplementary pensions, which supported the work of the current Minister of Pensions (Belgian Court of Audit, 2020^[37]). Nevertheless, the Court of Audit reports that its audits, albeit sent to Parliament, are rarely actually discussed by members of Parliament. Even when they are, there is rarely any impact beyond a hearing or discussion in committees – that is to say that the Court’s Recommendations are not often implemented.

Furthermore, Regulatory Impact Assessments could be implemented in a more efficient and effective way in order to fully exploit the potential of such instrument. One of the key issues underlying the implementation of RIA is the timing. These assessments are generally done at the very end of the policy-making process when the main policy options have already been decided. As a result, RIA often becomes a formalising tool for policies that were already acted and thus voiding RIA of its evaluative nature. The helpdesk provided by the Administrative Simplification Agency (ASA), which aims at improving the qualities of RIA, is rarely used. This is compounded by the fact that the ASA suffers from structural understaffing issues, which undermine their capacity to work to promote the quality of RIAs – and by extension the use of proper evaluations in the regulatory cycle.

A legal framework for evaluation across government helps to promote systematic linkages between supply and use of evaluations in decision making

While individual competencies are important, it is formal organisations and institutional mechanisms that lay the foundation for evidence-informed policymaking that can withstand transitions between leadership (Results for America, 2017^[35]). Institutional also enable the incorporation of policy evaluation findings into decision making by linking evaluations to the budget or policy cycles. By adopting a legal framework for policy evaluation, the Belgian federal government could clarify how evaluation findings are to be used for decision making. In the United States, for example, the Foundations for Evidence-Based Policy-making Act of 2018 requires agencies to use evidence whenever possible to further both mission and operations, and to commit to build evidence where it is lacking (115th Congress, 2019^[12]). This means, that government agencies in the US are expected to develop “processes and practices that establish habitual and routine reliance on evidence across agency functions and demand new or better evidence when it is needed.” Another interesting practice in this regard is that of the Canadian Policy on Results, which stipulates that evaluation results have to be submitted by department heads when they submit a proposal to the Treasury Board for new spending (Canada Treasury Board, 2016^[13]). Similarly, the legal framework at the federal level in Belgium could strengthen the link between policy evaluations and decision making.

Publicity and communication of results

Policy makers and stakeholders cannot use evidence and the results of evaluation if they do not know about it. The first step to promoting use is therefore that the results be made available to their intended users – simply put, that they be made public. But publicity alone is not enough, and active communication and dissemination strategies are needed to promote use.

There is no concerted and centralised effort to make evaluations public and to make them easy to access

Making evidence public is an important element to ensuring its impact: if citizens are aware of evidence, it may build pressure on policy makers to use it (OECD, 2020^[31]). Indeed, the publicity of evaluations is important in order to ensure that the public trusts the government not to “cherry-pick” evidence (OECD, 2017^[38]). Therefore, many countries have instituted policies mandating the publication of evaluations as well as to enhance the easy public access to these materials.

In the Belgium federal government, evaluations are not systematically published, which means that results are not always visible and easily accessible to users. Among Belgian federal services, only the Ministry of Social Integration and the Special Evaluation Office systematically publish evaluation reports on a dedicated web page (SPP Intégration Sociale, consulted 2018^[39]). In general, in-house evaluation reports are rarely made public and are rather communicated to the ministerial cabinet. According to the survey conducted by the Belgian Supreme Court of Audit, the reasons for such rare publications are confidentiality, financial reasons, and lack of approval for publication from the cabinet (Belgian Court of Audit, 2018^[5]).

Furthermore, certain ministries have been reducing the number of evaluations that they make public. For instance, the external academic evaluations commissioned by the Ministry of Interior, especially those related to prevention which were systematically published, have been stopped due to budget considerations. (Belgian Supreme Court of Audit, 2018). Indeed, the publication of internal evaluations requires significant time and financial resources, which ministries either lack or dedicate in priority to other activities.

Active communication of evaluation results across the federal administration is very limited

While a useful first step in promoting access to the evidence, publicity is not enough. Indeed, research suggests that in isolation, publicity alone does not significantly improve uptake of evaluations in policymaking. Rather, the presentation of evidence should be strategic and driven by the evaluation’s purpose and the information needs of intended users. As such, evaluation results ought to be well synthesised and tailored for specific users for their use to be facilitated.

A communication strategy should enable promoting the uptake of evaluation evidence by contributing to tailoring evaluation evidence to different publics. Indeed, for evaluation evidence to be used, it should not only be accessible to the public and policy makers but should also be presented in a strategic way and driven by the evaluation’s purpose and the information needs of intended users. None of these elements seem to be present in Belgium’s evaluation system, even though some individual ministries do present the results of their evaluations on a very *ad hoc* basis through seminars for instance.

A centralised database and an active communication strategy can increase the use of policy evaluations in the Belgian federal government

A first step in ensuring that evaluations are used in policymaking at the federal level in Belgium would be to ensure facilitated and standardised access to evaluations, either through the individual websites of the commissioning institutions or through a centralised database, such as the Norwegian evaluation portal (see Box 4.3).

Box 4.3. Centralised evaluations portal in Norway

In Norway, the Directorate for Financial Management and the National Library of Norway maintain and manage a centralised evaluations portal (<https://evalueringsportalen.no/>). All the studies and evaluations are made available on the portal as soon as published. Moreover, they are easily searchable and categorised. One can search based on topic, commissioning institution, conducting institution, type of evaluation (*ex post* evaluation, socio-economic analysis, etc.) or based on the underlying method of the study (based on questionnaires, public datasets, literature review). The portal contains the studies conducted since 2005 by the government and agencies as well as some selected earlier governmental studies. Finally, on the portal one can find various evaluation guidelines as well as evaluation agendas, relevant professional and news publications.

Such a centralised platform helps to build and enable the reuse of knowledge. Moreover, since it is easily searchable and updated by default it increases the transparency of public sector analysis.

Source: OECD (2020^[3]), *Improving Governance with Policy Evaluation: Lessons from Country Experience*, <https://doi.org/10.1787/89b1577d-en>.

Indeed, research suggests that ease of access is also an important factor in promoting use of evidence. The SPF BOSA could create an evaluation easy-to-use database that would make it easy to sort through the material based on the type of analytical material (e.g. evaluation of structural funds, regulatory impact assessment, *ex post* evaluation, etc.) and the institution that conducted the study. Publication of evaluations could also be made systematic and automated so that all the studies and evaluation can be found there as they are published. This repository or database could also link to (or house) regional or community level evaluations, to even further increase the impact of evaluations in decision making at all levels of government in Belgium.

Furthermore, use also requires active communication from evaluation suppliers. Thus, the SPF BOSA could develop a tailored communication and dissemination strategy, closely tied to the evaluation agenda or plan, in order to increase interest in evaluation findings. This communication strategy could include systematic use of executive summaries, tailored syntheses of research evidence, seminars to present research findings, as well as presentation of findings or data through social media, etc.

The SPF BOSA can position itself as a knowledge broker within the federal government

Taken as a whole, knowledge broker organisations tend to fulfil four main functions:

- **Generating knowledge:** They ensure that there is enough relevant evidence available for decision makers to answer pre-defined questions by synthesising the available evidence and by identifying knowledge gaps. If knowledge gaps are identified, knowledge brokers can fill them by commissioning research.
- **Synthesising knowledge:** They review existing evaluations to synthesis the evidence and identify potential gaps in knowledge.

- **Translation of knowledge:** When the required evaluation is available, they translate evidence gathered in a language that is understandable for decision makers.
- **Facilitating evidence adoption:** They build an organisational culture for effective adoption of evidence. They are involved in capacity building activities close to the public service providers and they build networks between knowledge producers and knowledge users. Additionally, they build and maintain informal relations with their stakeholders.

Whilst the SPF BOSA most likely does not have the ambition to be generating many evaluations itself, it has a key role to play within the federal government in synthesising existing evaluations (both national evaluations, but also from other countries or international organisations) in order to identify any gaps in information or, on the contrary, 'push' this information to the relevant user institutions. In the same way, the SPF BOSA may also, as discussed previously, play a key role in translating evaluations for users through active communication strategies, for instance, as well as build networks between evaluation producers and users.

Summary of recommendations

- Consider, along with an evaluation agenda or plan, increasing legislative planning to allow ministries to anticipate analytical needs over the course of a government mandate.
- Develop senior civil servants and policy makers' skills for use of evaluation. The SPF BOSA could organise training for Senior Civil Service leadership on these issues. Such trainings can take the form of workshops, masterclasses or seminars.
- By adopting a legal framework for policy evaluation, the Belgian federal government could clarify how evaluation findings are to be used for decision making.
- Create an evaluation easy-to-use database that would make it easy to sort through evaluations based on the type of analytical material (e.g. evaluation of structural funds, regulatory impact assessment, *ex post* evaluation, etc.) and the institution that conducted the study.
 - Systematically publish evaluations on this database.
 - Link to (or house) regional or community level evaluations.
- Develop a tailored communication and dissemination strategy, closely tied to the evaluation agenda or plan, in order to increase interest in evaluation findings. This communication strategy could include systematic use of executive summaries, tailored syntheses of research evidence, seminars to present research findings, as well as presentation of findings or data through social media, etc.
- Position SPF BOSA as a knowledge broker.

5 The role of BOSA in fostering a policy evaluation culture

Introduction

While BOSA's role has expanded with the creation of the Directorate General for Budget and Policy Evaluation, the challenge is to assess whether its formal mandate, structure and resources have sufficiently evolved to meet the new ambitions and needs associated with this shift in its role. In this context, this chapter offers an overview of BOSA's current mandates and suggests strengthening its activities in the area of standard setting for evaluation, promotion of use, co-ordination, etc. in order to increase its legitimacy and impact. Finally, the chapter identifies some operational challenges related to BOSA's human resources and organisation, and suggests that BOSA pursue its transformation to better reflect its new mandate.

Co-ordinating evaluations across the federal government

There is very little co-ordination across government on policy evaluation to date

Across the federal government, the different entities involved in evaluations are not acting in a concerted manner. Despite the existence of evaluation units in certain ministries, there is very little to no co-ordination between the different units. Even co-ordination within a single ministry can prove challenging between different evaluation units or teams across directorate generals.

This siloed approach means that administrations are much less likely to conduct cross-sectoral evaluations that cover policy issues that pertain to different ministries, as there is no impetus or incentive to collaborate on such issues. Nevertheless, policies are increasingly interconnected, particularly in a federal government where sub-federal entities are involved in all policy domains. This means that conducting *policy* evaluations, which look at complex public interventions comprising often a combination of legislations, regulations and programmes or projects, requires working across administrative siloes. The current organisation of the evaluation function across the federal government is therefore a missed opportunity. Such a silo-ed approach between ministries also means that evaluation resources may be wasted through duplication of efforts. Additionally, this model means that knowledge, best practices and skills stay at the level of individual ministries or teams, with few positive externalities across administrations.

The SPF BOSA has scope to become a key actor in the system

The Policy and Support Federal Public Service (SPF BOSA), has been given a role with respect to policy evaluation with the creation of the Directorate General for Budget and Policy evaluation. However, the exact scope of its mandate still needs to be defined more precisely at its organisational decree, the Royal

Decree of 22 February 2017 (National Government of Belgium, 2017_[20]), does not mention policy evaluation explicitly as being part of the SPF BOSA's mandate.

Ministries across the government see a clear need for more co-ordination of evaluation activities and have expressed their support for an increased role for the SPF BOSA in this regard. Indeed, the SPF BOSA is, in general, a well-respected actor and recent experiences, such as the establishment of the DG Digital Transformation as one of the key federal data players, has been seen as particularly helpful by line ministries. Through this example, the SPF BOSA has established its value added as the key player for inter-ministerial co-ordination in the field of strategic data governance.

Moreover, OECD data suggests that, in other a large majority of OECD countries, the actors that co-ordinate evaluations across government tend to typically lie in the Ministry of Finance, the Centre of Government and/or the Ministry of Public Sector Reform (see Figure 2.3). By combining some or all of the functions of these key ministerial portfolios, the SPF BOSA is well-placed to play this role.

The above-mentioned co-ordination institutions have key roles to play in the evaluation eco-system in many OECD countries, by making sure that evaluation can take place at the right time and in the right place and that it can feed into decision making. In some cases, the core institution can also develop its own capacity for evaluation (see Box 5.1).

Box 5.1. Range of responsibilities of policy evaluation co-ordinating institutions

In OECD countries, the main institution(s) in charge of co-ordinating evaluations across government carry-out typically carry out most or all of the following activities:

- Defining and updating the evaluation framework
- Developing guideline(s)
- Providing incentives for carrying out policy evaluations
- Undertaking policy evaluations
- Requiring government institutions to undertake specific policy evaluations
- Defining course of action for commissioning evaluations
- Developing skills, competences and/or qualifications of evaluators
- Developing standards for ethical conduct
- Ensuring quality standards of evaluations
- Promoting stakeholder engagement in evaluations
- Overseeing the evaluation calendar and reporting
- Promoting the use of evaluation
- Serving as a knowledge centre and providing a platform for exchange

Source: OECD (2020_[3]), *Improving Governance with Policy Evaluation: Lessons from Country Experience*, <https://doi.org/10.1787/89b1577d-en>.

Currently, the SPF BOSA aims to serve as a knowledge centre and platform for exchange, by fostering a network of evaluators in the public administration – although it can play this role to a very limited extent. This role would be quite welcome, especially in a context where there is no federal evaluation association. There is much scope for greater co-ordination from SPF BOSA and for this administration to take on additional responsibilities amongst those listed above.

In particular, SPF BOSA is well placed to address the skills and capacity gap found in ministries in regard to policy evaluation, especially considering that it also houses the Directorate General for Recruitment and Development. It could therefore take a leading role in nurturing a network of skilled analysts. For instance, SPF BOSA could foster and manage a network of analytical capacities across ministries and agencies. By giving seminars, sharing knowledge management and developing methodological guides for analysis and evaluation, SPF BOSA could also support the continuous development of public sector skills for evaluation.

Additionally, given the role SPF BOSA's new role in following up and reporting on the implementation of the national Resilience and Recovery Plan (RRP) (Federal government of Belgium, 2021^[40]), in addition to its role in the spending review process, it is well placed to play a more general role in the promotion of a results-based culture in the Belgian federal administration. The Resilience and Recovery Facility is a performance-based tool, which requires member states to gather performance data on the implementation of the reforms set out in their RRP. This data will be key to feed into policy evaluations, and, conversely, evaluative evidence can serve to assess the performance of ministries in implementing the plan. As a key player in the RRF system and in co-ordinating spending reviews, the SPF BOSA can leverage this new role to push for more evaluations in line ministries and to review existing performance evidence for the purpose of following up on the RRP reforms. As for spending reviews, given the similarities as well as the interaction between policy evaluations and spending reviews (they should complement each and can be used as input for one another), a co-ordinated approach within BOSA for both tools would be relevant.

A resources and structure

In spite of the SPF BOSA's new role in regard to policy evaluation, and the subsequent creation of the DG Budget and Policy Evaluation, the service's resources and structures have not been appropriately updated to match these new responsibilities. Indeed, there is currently only one person in BOSA working on policy evaluation, not even on a full-time basis. These are clearly not sufficient capacities in order to play an important role in the federal government evaluation system, even if this role does not involve carrying out many evaluations itself. For SPF BOSA to establish itself as a credible player, it will need to have the appropriate resources and skills to be in a position to serve as a resource for other ministries seeking guidance on evaluation.

Summary of Recommendations

- Give a formal mandate to SPF BOSA in the area of policy evaluation by updating its organisational decree. In particular, SPF BOSA could:
 - Develop general quality guidelines for policy evaluation.
 - Develop, in co-operation with the line ministries, an evaluation agenda or plan.
 - Co-ordinate complex cross-sectoral evaluations and analyses across ministries.
 - Engage with a community of evaluators across ministries, sharing methods, organising seminars and peer review of the work.
- SPF BOSA could further help address the analytical capacity and skills gaps within the Belgian public sector through:
 - Managing the annual recruitment and the selection of a set of professional analysts for the government.
- Align the resources and internal organisation of SPF BOSA in order to ensure that it can play this role.

References

- 115th Congress (2019), *Public Law No: 115-435 (01/14/2019) - Foundations for EvidenceBased Policymaking Act of 2018.* [12]
- Belgian Court of Audit (2021), *First Recruitments. Target Group Reduction of Employer Social Security Contributions.* [36]
- Belgian Court of Audit (2021), *Presentation on the role of the Court to the OECD.* [17]
- Belgian Court of Audit (2020), *Supplementary pensions - Efficiency of public policy on social and tax incentive..* [37]
- Belgian Court of Audit (2018), *Ability of the federal government departments to assess public policies.* [5]
- Belgian Court of Audit (1998), *Loi modifiant la loi du 29 octobre 1846 organique de la Cour des comptes.* [16]
- Canada Treasury Board (2016), *Policy on Results*, <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=31300> (accessed on June 2022). [13]
- Federal government of Belgium (2021), *Plan National pour la reprise et la résilience.* [40]
- Federal government of Belgium (2020), *Accord de gouvernement.* [4]
- Fobé, E., Peuter, B., Jean, M., & Pattyn, V. (2017), *Analytical techniques in Belgian policy analysis.*, Bristol University Press. [24]
- France Stratégie (2016), *Comment évaluer l'impact des politiques publiques: un guide à l'usage des décideurs et des praticiens.* [32]
- INTOSAI (2019), *Guidelines for Internal Control Standards for the Public Sector.* [10]
- Ministry of Finance of the Netherlands (2018), *Arrangements for periodic evaluation research*, <https://wetten.overheid.nl/BWBR0040754/2018-03-27> (accessed on 2022). [11]
- National Government of Belgium (2017), “Arrêté royal du 22 février 2017 portant création du Service public fédéral Stratégie et Appui”. [20]
- National Government of Belgium (2014), “Arrêté royal modifiant l’arrêté royal du 25 février 2010 portant création d’un service Evaluation spéciale de la Coopération internationale”. [15]
- National Government of Belgium (2013), *Arrêté Royal fixant le règlement organique du Service public fédéral Finances ainsi que les dispositions particulières applicables aux agents statutaires.* [14]

- National Government of Belgium (2013), *Loi portant des dispositions diverses concernant la simplification administrative*. [6]
- National Government of Belgium (1994), "Loi du 21 décembre 1994 portant des dispositions sociales et diverses". [19]
- OECD (2022), "Best Practices on Spending Reviews", *OECD Journal of Budgeting*, <https://doi.org/10.1787/16812336>. [8]
- OECD (2022), *Building Trust to Reinforce Democracy*. [1]
- OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", *OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, <https://doi.org/10.1787/483507d6-en>. [21]
- OECD (2022), *Recommendation of the Council on Public Policy Evaluation*. [2]
- OECD (2021), *Applying Evaluation Criteria Thoughtfully*, OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>. [33]
- OECD (2021), *Mobilising Evidence at the Centre of Government in Lithuania: Strengthening Decision Making and Policy Evaluation for Long-term Development*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/323e3500-en>. [22]
- OECD (2021), *OECD Regulatory Policy Outlook 2021*, OECD Publishing, Paris, <https://doi.org/10.1787/38b0fdb1-en>. [7]
- OECD (2021), "The OECD Framework for digital talent and skills in the public sector", *OECD Working Papers on Public Governance*, No. 45, OECD Publishing, Paris, <https://doi.org/10.1787/4e7c3f58-en>. [28]
- OECD (2020), *Building Capacity for Evidence-Informed Policy-Making: Lessons from Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/86331250-en>. [34]
- 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>. [3]
- OECD (2020), *The Irish Government Economic and Evaluation Service: Using Evidence-Informed Policy Making to Improve Performance*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/cdda3cb0-en>. [25]
- OECD (2020), *The OECD Open useful reusable data (ourdata) index*, OECD, Paris. [26]
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [30]
- OECD (2017), *Policy Advisory Systems: Supporting Good Governance and Sound Public Decision Making*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/9789264283664-en>. [38]
- OECD (2017), *Skills for a High Performing Civil Service*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/9789264280724-en>. [23]

- OECD (2011), *Internal Control and Internal Audit: Ensuring Public Sector Integrity and Accountability*, OECD, Paris, <https://www.oecd.org/governance/47638204.pdf>. [9]
- Results for America (2017), *Government Mechanisms to Advance the Use of Data and Evidence in Policymaking: A Landscape Review*. [35]
- SPP Intégration Sociale (consulted 2018), *Études, publications et statistiques*, <http://www.mis.be/fr/etudes-publications-statistiques> (accessed on 15 February 2023). [39]
- Statbel (2022), *Microdata for research*, <https://statbel.fgov.be/en/about-statbel/what-we-do/microdata-research> (accessed on 15 February 2023). [27]
- Stockmann, R., W. Meyer and L. Taube (eds.) (2020), *The Institutionalisation of Evaluation in Europe*, Springer International Publishing, Cham, <https://doi.org/10.1007/978-3-030-32284-7>. [18]
- van Ooijen, C., B. Ubaldi and B. Welby (2019), “A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance”, *OECD Working Papers on Public Governance*, No. 33, OECD Publishing, Paris, <https://doi.org/10.1787/09ab162c-en>. [29]
- World Bank Group (2019), *World Bank Group Evaluation Principles*. [31]