

## Chapter 5

### Belgium<sup>1</sup>

*Valérie Pattyn and Bart De Peuter*

#### 1 General country overview

In order to have an accurate understanding of the level and way of institutionalisation of policy evaluation in Belgium, it is essential to highlight some key tenets of its politico-administrative setting: the federal structure, multiple party coalition governments and partitocracy.

Since 1993, Belgium is officially a *federal state*. Consecutive constitutional reforms have transferred major competences from the national (coined ‘federal’) level to the regional level. The distribution of competences on the regional level is mainly organised along two broad axes. The first axis concerns linguistics and culture. A Flemish-, French- and German-speaking Community are each in charge of person-related matters (examples: education, culture, youth). The second axis is designed to accommodate the diversity in socioeconomic interests in the country and has constituted the basis for the creation of three territorial regions: the Flemish Region, the Walloon Region and the Brussels-Capital Region. The regions are in charge of territorial matters (examples: environment, agriculture), and legislation with regard to municipalities and provinces. Despite the devolution in competences to regions and communities, the federal level in Belgium is still competent for traditionally core tasks of a state including internal security, foreign affairs, defence, justice and finance. Moreover, in several domains, competences are scattered over the federal and regional level (examples: labour market, health care and social security). All government levels do have their own government; parliament and administration, with the exception of the Flemish region and community which have merged their institutions (Pattyn and Brans 2015).

---

<sup>1</sup> This chapter is to large extent based on the conference paper entitled ‘*Once Laggard, Always Laggard? A multi-level analysis of the state of evaluation institutionalisation in Belgium*’, that the authors presented at the 2016 European Evaluation Society conference in Maastricht. For this conference paper, we launched a questionnaire that was modelled along the survey used in (Jacob et al., 2015). We owe special thanks to Steve Jacob for sharing this questionnaire; and we are most grateful to the eleven respondents that provided us with valuable input on the state of evaluation institutionalisation in Belgium. The insights collected via this survey have been integrated in this chapter.

Secondly, and related to the consensus seeking culture in Belgium, the political landscape is characterised by a system of proportional elections. This, in turn, triggers a fragmented party-political landscape and *multiple party coalition governments* on all levels. To be clear, since the 1970s political parties are also regionalised in their organisation, with the exception of the extreme left-wing Workers' Party (PvdA/PTB).

*Partitocracy* constitutes a third key characteristic of the Belgian politico-administrative setting (Deschouwer, 2012; Dewachter, 2001). In a country that is highly fragmented, political parties act as gate keepers for demands coming from the executive and legislative branches and play a decisive role in the policy-making process. Linked to the partitocracy, is the influential role of personal staffs of ministers which are large by international standards and dominant in policy formulation (Brans et al. 2005). The advisors assist the minister in identifying problems, in outlining policy and in everyday decision-making. They also play a pivotal role in cross-departmental coordination.

Why are these country features important to the study of evaluation institutionalisation in Belgium? First, the complexity of the federal structure makes it challenging to describe its state of evaluation praxis in a manner that gives full justice to the different administrative cultures and evolutions on the different governmental levels that exist in the country. Belgium has not been included in reputed publications that describe the level of evaluation maturity and institutionalisation across OECD countries (including Furubo et al. 2002; Jacob et al. 2015). An exception is the study of Varone, Jacob and De Winter (2005). Secondly, for decades the process of federalisation has required a lot of political and administrative energy and focus, with an impact on attention given to and the degree of structural policy analytical capacity building.

As a matter of fact, the partitocracy and the influential role of the large ministerial cabinets intensify the area of tension between a political rationale and a more rationalistic perspective to which evaluations can contribute. As Varone, Jacob and De Winter (2005, p. 262) have put it: "As evaluators can call into question policy choices, which are the products of delicate compromises between four to six coalition parties, it is quite obvious that the party leaders would resist this potential loss of coordinating power".

Against these backgrounds and from the few and earlier available comparative studies (Jacob 2004; Jacob 2005; Varone et al. 2005) Belgium is to be considered as a latecomer in terms of evaluation culture and praxis to be situated in the slipstream of the second wave with which evaluation praxis has spread in Europe during the early 1990s. Rather external pressures have put evaluation on the agenda, such as evaluation demands and impulses from the EU (Furubo et al. 2002). In addition, New Public Management (NPM)-inspired government wide administrative reforms have played a major role. Yet, intra-country differences exist:

evaluation demands have reached Wallonia more and earlier than Flanders. The explanation lies mainly in the evaluation requirements related to the European co-financed programmes (i.e. structural funds). Irrespective of these intra-country differences, from an international comparative perspective, the important conclusion anno 2004-2005 was that Belgium, after a late start in the 1990s, had to be classified as a laggard in terms of evaluation institutionalisation (Jacob 2005; Jacob, 2004; Varone et al. 2005).

In this chapter we take stock of the recent developments in policy evaluation praxis more than one decade later. We describe the state of institutionalisation of evaluation in Belgium with a view to the federal and regional level, the latter with some bias to the Flemish region.

## **2 Institutional structures and processes (political system)**

### **2.1 Evaluation regulations**

#### **2.1.1 A formal anchorage of evaluation requirements in laws or regulations**

In Belgium the practice or use of policy evaluation as such is not anchored in the constitution, or / neither in national laws or regulations. Nor does a central country-wide policy or strategy on policy evaluation exist. The above-described country-specific features of the political system have contributed to the spread of the perception that evaluation is functional for accountability, rather than for policy learning or improvement.

However, in the last 20 years policy evaluation has been put on the agenda across the entire spectrum of policy domains. At the same time, government-specific laws and regulations have been introduced on the implementation of generic or sector-specific evaluation tools such as regulatory impact analysis (RIA) and environmental impact analysis (EIA).

Following supranational regulation (EU) and advocacy (OECD), the implementation and use of some evaluation tools with broad scope have been regulated in Belgian domestic law, both at the federal and at regional level.

At the federal level, regulatory impact assessment is anchored in the decision-making process by law. It is mandatory for all draft legislation, ministerial and royal decrees, excluding draft legislation constituting the ratification of international treaties, intergovernmental agreements or auto-regulation regarding the federal government. In an effort to better coordinate the existent federal thematic ex ante tests, it was decided in 2013 to integrate a number of them about topics such as administrative simplification (established in 2004), sustainable development (2007), gender (2007), Small and medium-sized Enterprises (SMEs 2008) and policy coherence for development aid (2013) in a more generic RIA framework (DAV 2014). It remains to be seen whether the

integrated RIA obtains its goal of raising awareness for assessing the impact of policy decisions, and whether it will overcome difficulties which the former separate tests encountered, that is limitations to self-evaluation and poor-quality control from the overseeing agency (OECD 2015).

In Wallonia, a similar integrated procedure for *ex ante* evaluation of proposals of legislation exists, but the Walloon regional government refers to this set of tests as ‘an examination of conformity of sustainable development goals’ (Regional Law of 27 June 2013). The Walloon version can be considered as a light variant of the federal RIA. The appraisal is carried out by a central unit within the government and its result is formulated as an internal advice to the government.

In Flanders, the introduction of RIA was regulated in 2000 as part of a reform package for better regulation. Since 2005, the RIA is required for all drafts of parliamentary decrees and governmental acts which have a regulatory effect on citizens, private or non-profit organisations. As of now, the RIA has developed into a standard procedure in policy analytical work, specifically as regards the preparation of policy initiatives that are translated into regulation. It intends to outline the pros and cons of possible policy options by comparing expected direct and indirect effects.

All in all, the relevance and impact of the RIA should not be overestimated. While it is meant as a tool supporting policy-makers in their decisions, it quite often serves as an after the fact justification for decisions that have arrived through political bargaining (Fobé et al. 2017b).

In addition to the RIA, the well-known EU developed EIA has been converted in regional regulation with specific accents. Although linked to environment policy, this *ex ante* evaluation technique is used for a wide range of governmental decision-making. In Flanders, the Decree on environmental impact assessment of 2002 (amending the Decree of 1995 containing general regulations on environmental policy) translated the European Directives, and adapted it to the Flemish context. The Flemish variant includes an additional participatory stage in the scoping phase (see below). The administrative regulations further require a systematic and scientific analysis of expected environmental impacts of the intended plan or project and of its ‘reasonably to be considered’ alternatives. In addition, the assessment report should have a conclusion in which the best acceptable alternative and recommendations are outlined, in terms of mitigating measures to avoid, limit, remedy or compensate for the substantial negative effects.

Government-wide public sector reform initiatives have also been a major catalyst for the adoption of evaluation practice and use, especially in Flanders. Inspired by NPM, the so-called ‘better administrative policy’ reform in 2003 was aimed at a decoupling of the policy cycle. Departments were to be conceived as core

administrations responsible for planning and evaluation, and autotomised agencies were assigned the responsibilities for policy implementation. Departments were also charged with the explicit obligation to produce outcome evaluations on the level of the policy domain. While the philosophy of the reform was meant to concentrate evaluation capacity in the departments, the present-day situation, more than a decade later, shows considerable variation when comparing the different policy domains (see below). Also, the overall outcome evaluations have never been developed. Nevertheless, the agenda-setting role of the reform for evaluation must not be underestimated.

Hence, despite government-wide initiatives, sectoral dynamics still prevail and have resulted in substantial variation across policy domains. The latter also holds for sectoral regulations and policies on evaluation. In many policy fields, the institutionalisation of evaluation has been strengthened in the past 10 years. Indications for this—apart from increased practice—can be found in the presence of an evaluation framework, a long term planning of evaluation activities and/or systems of evaluation quality control (Desmedt et al. 2016). An explicit recognition of the evaluation functions, also in organisational charts, proved to be an important leverage in this regard.

By way of illustration, we concisely elaborate on the institutionalisation of evaluation practice within a selection of policy sectors at the Flemish regional level (environment; education; and economics, science and innovation), and at the federal level (development cooperation).

Environment has traditionally been a leading sector when it comes to policy evaluation in Flanders. In 1995 and 1997, services within government agencies were given the duty to publish evaluation reports. This was deemed compatible with an integrated approach towards policy planning, reporting and annual programming. A lack of expertise in evaluation, however, delayed the genuine establishment of the policy evaluation pillar until the millennium switch. Although considerable evaluation activity was developed at that time, it appeared not to be sustainable despite the existence of the above-mentioned regulations and evaluation clauses in environmental legislation (see below).

A more sustainable development can be found in the field of education, another domain with a relatively long evaluation tradition and high evaluation activity. In the education field, one can identify several requirements for evaluation in various thematic decrees; yet there is no overarching regulation on evaluation. Nonetheless, the administration of the education sector invested considerably in evaluation practice (quality) and capacity via the establishment of a separate unit which coordinates and supports evaluation activity. Exemplary are the introduction of a tailor-made evaluability assessment tool by this unit, and the development of an introductory

manual that is widely diffused within the educational administration.

Similar dynamics can be observed in ‘economics, science and innovation’ where a specific departmental unit is in charge of the development of an evaluation agenda and of ensuring evaluation quality control in its area.

At the federal level, we would like to highlight the field of international development as belonging to the high-end policy sectors in terms of evaluation institutionalisation and as having a strong regulatory framework. The Special Evaluation Office for Development Cooperation was founded in 2003 by Royal Decree, and was updated in 2010 and 2014. The Office is assigned the tasks of planning, conducting and follow-up of evaluation regarding every form of development cooperation (co-) financed by the federal state. Its evaluation focus lies on relevance, efficiency, impact and sustainability of interventions. The tasks of the service were further translated in a policy document in 2014. The framework stresses the need for cooperation with all parties involved in the planning and application of evaluation processes in order to match the stakeholders’ needs, to enhance the credibility of evaluations and to ensure that findings and recommendations trigger improvements and new dynamics in both policy and practice. It is exactly the combination of a sound legal basis and the strategic and operational elaboration in a policy document that constitutes a firm basis for the Office’s evaluation practice (De Maesschalck et al. 2016).

### **2.1.2 The institutionalisation of evaluation within parliament**

In principle, one could view Parliament as an important trigger of evaluation demand, as it regularly asks for the inclusion of evaluation clauses in legislation. In fact, in the neo-corporatist and consensus-style policy context of Belgium, these clauses are often used as a mechanism to overcome a political deadlock and function as leverage to achieve a compromise in difficult negotiations over policies. On the federal level the insertion of evaluation clauses is rather limited. In the Flemish Parliament the management services keep an inventory of all reporting clauses in decrees; the numbers are revealing. Between 1971 and 2013 in total 285 clauses were identified, with differences in frequency (once, yearly, other periodicity) and level of concreteness. About 38 concerned an explicit obligation to conduct an evaluation and to deliver it to Parliament. Yet, in the absence of rigid follow-up mechanisms, only in six out of the 38 cases an evaluation was actually conducted (SERV 2015).

When compared to the executive branch of government, the institutionalisation of evaluation within the legislative branch is clearly less present. Hindrances are the lack of specific structures and the absence of evaluation-specific budgets. At the federal level, a parliamentary committee for legislative evaluation was founded and has been operational between 2011 and 2014. Yet, it now only exists on paper. Nevertheless, the

balance is not merely negative. Several attempts toward a stronger embedding of evaluation in parliamentary praxis can be identified. In the Flemish Parliament, under impulse of its chairman, a thorough reflection operation was recently organised along different scenarios for a more systematic ex post decree evaluation. The Social-Economic Council suggested a better follow-up of evaluation clauses and the inclusion of more evaluative information in yearly policy letters. In addition, it called for a broader use of existing instruments such as thematic debates and hearings, a stronger cooperation with the Court of Audit and the need to draft an evaluation agenda (SERV 2015). It remains to be seen, however, whether this will evolve toward stronger institutionalisation.

Besides this, evaluations are commissioned or conducted by Parliament in the context of inquiries triggered by exposed (indications of) public sector dysfunctions (e.g. railway safety policy after train accident, anti-terrorism policy after terrorist attacks) or with a thematic focus (e.g. national commission for the evaluation of the law on abortion). Reports of the Belgian Court of Audit are sometimes—not always—discussed. The utilisation or reports usually ends after the hearings, though.

Against the low level of institutionalisation in Parliament, one should acknowledge that a substantial number of members of parliament does show attention and interest for evaluation, as an analysis of parliamentary written questions in the Flemish parliament revealed (Speer et al. 2015). The initiation, the results and the use of evaluations turned out to be of most interest to parliamentarians. Interestingly, both majority and opposition party members frequently ask evaluation related questions.

For Wallonia, observers noticed that the Walloon Institute for Evaluation, Foresight and Statistics (IWEPS) is less frequently demanded for input by Parliament. IWEPS is the central unit for evaluations commissioned by the Walloon Region (see below). Also, for this region and in the Brussels Capital region, legislation often contains evaluative clauses leading to evaluation activities. Yet, specific structures or procedures for the dissemination and use of evaluation findings do not exist.

## **2.2 Evaluation practice**

### **2.2.1 A relatively wide diffusion of evaluation practice**

As we mentioned in the introduction, Belgium was a late starter, to be situated in the slipstream of the second wave of countries in which evaluation has been introduced in the late 1990s (Furubo et al. 2002). Policy evaluation activity is nowadays widely diffused across the entire public sector. Generally speaking, evaluation praxis (apart from RIA) seems to have developed more systematically on the regional level than on the federal level. EU demands for evaluation and public-sector reforms encompassing attention to evaluation have had a

clear catalyst function for the regional level. For the federal government level, such major catalysing stimuli cannot be identified.

Empirical indications for evaluation activity can be found in a large variety of sources, such as policy (planning) documents referring to existing or planned evaluation procedures and activities, announcements of public tenders and evaluation reports published on governmental websites, and via references to evaluation in the recent government agreements.

The current Flemish government agreement 2014–2019, for example, includes 50 explicit references to evaluation commitments. This number is but a small share of the total number of announced evaluations. Sectoral ministers can highlight additional evaluations in their so-called policy notes, that is formal communications in which ministers outline their policy intentions for the upcoming government term. In a longitudinal 20 year timespan analysis of references to evaluation, respectively 247, 254, 347 and 231 references (with context information on the purpose, cf. *infra*) were counted for the five-year terms between 1999 and 2019 (De Peuter and Pattyn 2016).

Substantial differences exist again across policy domains. Belgium largely follows international trends: the domains that are front running in evaluation in an international context are also the domains where we can detect most evaluation activity in Belgium; and vice versa, this also applies to laggard domains. Policy domains as education, employment, public health and development cooperation have a strong evaluation culture. Although policy evaluation in Belgium is not restricted to those areas, these are clearly the fields that distinguish themselves in terms of the frequency of evaluating (Pattyn 2015). The same can be said for social affairs at the Federal level, for economics, social policies and poverty in Wallonia and for the combined domain of economics, science and innovation, and for environment in Flanders. But as stated above the enthusiasm for evaluation of environmental policy has, except for the obligatory environmental impact assessments, faded away in recent years, mainly due to a lack of political commitment, an insufficient match between supply and demand concerning focus and too little coordination capacity within the department (Crabbé and Vansteertegem 2016). For Wallonia, we already mentioned the important role of IWEPS. As of 2004, the Walloon government has systematically charged this economic research institute with the evaluations of the successive Marshall Plans, that is large-scale economic redeployment programmes. Given the size of these programmes and their importance for Wallonia, the evaluation activity in this region for a large part relates to the domains of economic, social and employment policies. In 2015 IWEPS also evaluated the government's anti-poverty



programme. At the other side of the continuum we can situate policy domains such as justice and home affairs (federal level) or culture (regional level) where evaluation is predominantly taking place on an ad hoc basis.

One should be careful, of course, in classifying entire policy domains in terms of evaluation activity. Also, within certain policy domains, large heterogeneity often exists at the level of single public sector organisations. Even in leading fields, such as education, not all public sector organisations are equally active in policy evaluations and in policy domains that are generally not very active in policy evaluation, there are sometimes organisational islands with a strong evaluation culture (Pattyn 2014).

In 2015, the Research Centre of the Flemish Government launched a survey among leading civil servants of all policy sectors, across departments and agencies (population: n=368; response rate: n=186). The survey provides relevant insights as for the main initiators of evaluation: 64% of the respondents stated that the own division mostly or always takes the initiative of setting-up an evaluation. Half of them (50%) (also) pointed at a request from the minister, 30% to obligations in regulation and 20% referred to stakeholders asking for evaluations or evaluations following specific topicalities. Parliament was indicated by 16% as the most common or exclusive source of demand for evaluation (Verlet et al. 2015).

### **2.2.2 Type of evaluation practice**

A taxonomy of evaluation can use the timing (ex ante, interim, ex post) or the focus (input, process output, outcome) as distinguishing criterion.

When it comes to the *timing* of the conducted (ad hoc) evaluations, there is a strong bias towards retrospective (interim/ex post) evaluations, compared to ex ante evaluations. The embedded RIA applications are herein not taken into account.

When looking at the *focus* of evaluations, the different government levels have a relatively similar profile. For the federal level, evaluations with a focus on outcome do take place, yet output and process evaluations are comparably more numerous. Considerable monitoring and evaluation capacity is occupied with the follow-up of timely service delivery and the number of inspections. Less attention is paid to the effectiveness of the many measures that are mentioned in action plans. When also considering the ex ante impact assessments related to the introduction of new regulations, the picture is more nuanced. In the latter type of assessments, the focus lays on expected effects and impact of various policy alternatives.

As for Flanders, the above-stated survey results are telling (Verlet et al. 2015): 90% of respondents indicated that their evaluations ‘always’ or ‘almost always’ concern outputs; and 62% reported to ‘always to almost always’ investigate policy effects. Although survey results should be treated with care, it is clear that a

lot of investment and attention goes to monitoring activity and output evaluation. The latter is not always accompanied with effect or impact evaluations.

In the Walloon region similar trends can be identified. A lot of monitoring practice and output evaluations are being conducted. According to the experts we contacted, the outcome-focused evaluation of the economic Marshall plan, coordinated by IWEPS, is rather an exception.

### **2.2.3 Locus of evaluation practice**

As in-house policy evaluations are not always officially classified as ‘evaluations’, accurate data on the relation between evaluations that are conducted in-house and outsourced evaluations are not available. Research in the Flemish context (Pattyn and Brans 2013) concluded that most evaluations are conducted in-house, although there are departments (e.g. education) that systematically opt for external evaluations. For the other government levels, we expect to find similar trends.

Even when an administration decides to outsource evaluations, this is not always evident. Experts at all government levels voiced concern over the limited pool of external evaluators. The relatively small supply side on the Belgian evaluation market influences the extent of outsourcing evaluation. Moreover, as only a limited number of applicants usually respond to evaluation calls, inevitably the same evaluators are repeatedly evaluating the same policies time and again. Similarly, a handful of private evaluation firms can easily build competitive advantages in a limited market. Methodologically, the pool of techniques used is also restricted. Furthermore, peer reviews are hard to implement, as everybody knows one another (Pattyn and Brans 2013).

As to future trends, capacity plays a key role. In Flanders the available survey (Verlet et al. 2015) highlights a divided perception: 30% of the respondents reports to have sufficient internal capacity to meet the demands of evaluation addressed to their administrative entity, while about 33% experiences a lack of capacity. About 60%, in turn, states to have sufficient knowledge to conduct or steer external evaluations, yet only 30% feel that there is sufficient external capacity for evaluation. In terms of funding, an important share of 25% reports to have a dedicated budget for evaluation. Relevant to mention is the significant reshuffling of the Policy Research Centres (*Steunpunten*). In 2001, these Centres were established with the aim to conduct short-term applied research, as well as fundamental long-term basic research on negotiated themes for the government. Such consortia of university institutions existed in most policy domains. In the past they also conducted a considerable share of evaluative research. In 2014 the system of co-financing by the Science department was abolished though, resulting in fewer and smaller centres and leaving other policy domains with a choice between decentralised outsourcing, more in-house evaluation, or reducing the amount of evaluation activity.

#### **2.2.4 Evaluation institutes, with varying levels of independency**

During the previous decade, all government levels have taken major steps to institutionalise evaluation in the executive, through the creation of divisions or units that are (also) responsible to support the strengthening and harmonisation of evaluation practices.

As to the horizontal institutionalisation of evaluation practices at the federal level, the Administrative Simplification Division, situated within the office of the Prime Minister, played a major role with regard to the coordination of the introduction of the Regulatory Impact Analyses (RIA) across federal policy departments. Today, it hosts the secretariat of the Impact Analysis Committee, which provides advice and feedback on the implementation of the RIA with the aim of strengthening the quality of applications of this ex ante evaluation tool and legislation in general (Fobé et al. 2017b).

At the regional level, IWEPS has a prominent function in the horizontal institutionalisation of policy evaluation with regard to policy domains for which the Walloon Region is competent. It was explicitly charged with the task to centralise studies commissioned by the Walloon Region and act as a support agency regarding methods of analysis and policy decision (Fobé et al. 2017b). Within the Brussels-Capital Region the Institute for Statistics and Analysis performs evaluations of regional policies for its government since 2013. Within the French Community and in Flanders evaluation capacity is less coordinated. The Study Centre of the Flemish Government has been much more active on the support and coordination of monitoring instruments compared to evaluation.

Despite the existence of horizontal evaluation units, evaluation capacity development to large extent evolves along dynamics at the organisational and sectoral level. In organisational terms, policy domains predominantly function as disconnected silos. For policy evaluation, this implies that administrations are searching answers to the same issues but do not systematically exchange evaluation related knowledge, experience or documentation. The 'island-culture' also hampers cross-sectoral evaluations that concern transversal policy issues (such as sustainable development).

Across the public sector, situated either within the centralised departments or within the public agencies that operate at arm's length of these departments, units were created or given the additional task to conduct or coordinate evaluation. These have served as an important stimulus for the development of evaluation capacity in many public sector organisations (Pattyn 2015). Some actively conduct evaluation studies themselves (e.g. within the Flemish Employment Agency, the federal Service for Sustainable Development, and

the Research administration in the Walloon region). Others predominantly outsource evaluations and still other departments rely more on the input of horizontal services (Fobé et al. 2017b).

Although these units are visible in the organisational chart, it would require in depth research to accurately assess their actual extent of autonomy. In principle, most of them are not conceived to work on a fully independent basis. The earlier mentioned Special Evaluation Office for Development Cooperation is an exception. It has a large degree of autonomy, both legally vis-à-vis the Directorate-General for Development Cooperation and Humanitarian Aid and functionally in dealing with the planning, implementation and steering of evaluations. The Office is accountable to the public and the parliament, to which it submits an annual activity report (De Maesschalck et al. 2016).

Despite the formal institutionalisation of evaluation units at organisational level, observers have expressed concerns about the sustainability of the evaluation capacity in many organisations (Desmedt et al. 2016). The evaluation function is often performed by one or only a handful of civil servants which are strong believers in evaluation. These ‘evaluation entrepreneurs’ frequently act as the de facto evaluation memory in the organisation, in absence of a sound archive of evaluation practices. This reliance on a few experts constitutes a risk to evaluation capacity building in the longer run.

Also, the Belgian Court of Audit undertakes evaluative activities. Since 1998, the supreme audit institution has been given a mandate to check public policies on economy, efficiency and effectiveness, apart from its classic praxis with financial audits and legality and regularity audits. Since 1999 it conducts so-called performance audits (about 10 per year on average based on an internal count). While in the past this type of audit rather focused on the presence of preconditions for effective and efficient policy and on goal attainment, recent examples demonstrate a shift toward policy evaluation with a substantial judgement on outcome and impact of policy initiatives. In its strategic plan (2015–2019) the Court of Audit underlined the intention to endorse this shift with its goal to assess the policy results of societal relevant themes. Currently it tries to elaborate its evaluation capacity accordingly. It has full autonomy to select policies and programmes and does so based on a risk analysis and indications of potential for improvements. Moreover, as a collateral body of parliament it aspires to be a reference for parliamentary work, policy-making and public debate. The Court of Audit also fine-tunes its activities with the Flemish Audit Office. The latter acts as an internal audit service for the Flemish governmental administration with a focus on organisational control. The Belgian Court of Audit is competent for all governmental levels except for the local level (for which the regions can perform external audits of the local internal control systems). It must be noted, though, that the performance audits conducted by

the Court of Audit only count for a very small share in the total evaluation praxis in Belgium and its regions. The exact share depends on the very definition of policy evaluation, of course. Nevertheless, the institution's reports gain considerable attention from policy-makers. Follow-up of its audits is also monitored. Within the Flemish region for example, ministers have to explain to Parliament in their annual policy notes how they deal with recommendations received from the Court of Audit.

## **2.3 Use of evaluations**

### **2.3.1 Evaluation motives**

As Chelimsky (1997) argued: "The purpose of an evaluation conditions the use that can be expected of it" (Chelimsky 1997, p. 32). The other way around, the study of use of evaluation can start by looking at evaluation purposes. Given the lack of an explicit evaluation agenda, ministerial policy notes constitute a relevant information source in this regard. Reference to evaluation is often made when policy initiatives are announced. A longitudinal analysis of ministerial policy notes of the Flemish government in Belgium yielded interesting insights on intended use of evaluations (De Peuter and Pattyn 2016). The focus of the study were the evaluation purposes being mentioned in ministerial policy notes, issued in four subsequent government terms. In total, the ministerial documents span a period of 20 years of policy making, ranging from 1999 to 2019. We distinguished between three main groups of evaluation purposes: evaluations can be set up to underpin policy planning, they can serve accountability purposes, and/or they can provide input for policy improvement and policy learning. We highlight a few key observations. First of all, the total number of references across ca. 25 policy notes remained relatively stable for three out of four legislatures. Also, the distribution over the purposes did not show substantial variation, that is one third of the references relates to the purpose of policy planning, while almost two thirds refer to policy improvement and learning. The various policy domains overall follow the same pattern, in terms of this distribution. Differences are more outspoken when we compare the number of evaluation references across individual policy notes; some contain dozens, other only state a few. The 'champions' are again those sectors which are known to be frontrunners in evaluation in Belgium. Across all four legislatures, the domains of education, environment, economics, wellbeing, and traffic mobility make up for almost half of all references.

Of course, purposes of evaluation do not yet explain whether evaluation results are also used in the same way. As for actual use, little systematic evidence is available. But it is clear that policy sectors do differ in terms of the systematic use of their evaluations. Evolutions are also observable throughout time. For example, we already mentioned that evaluation activity in the field of environment dropped due to problems with

matching supply and demand, which does not contribute to the use of findings. Within the field of education, instead, the potential for evaluation use seems to benefit from the systematic participatory approach towards evaluations: stakeholders are actively engaged in the evaluation process. The field of international development at the federal level follows a similar participatory approach, by engaging guidance committees in the evaluation process. To be clear, the guidance committees are not implied in the preparation of the evaluation, nor in the follow-up of evaluation results. These roles were assigned to an advisory committee, but its involvement decreased after some time. In absence of systematic research, one can but speculate about the use of evaluation reports beyond the civil service. In part II of this chapter we provide insights in use by civil society organisations.

### **2.3.2 Quality assurance**

In the Flemish administration, quality assurance at an organisational level is overall rather limited, and mainly ad hoc. In a study (Pattyn and Brans 2014) among 18 organisations (departments and agencies) from a mixture of policy fields we screened for the systematic application of a range of common quality assurance measures. It turned out that the practice of involving second readers for evaluation reports is most spread. Common practice is further the establishment of a steering committee or working group for evaluations. Evaluation manuals are neither common practice, with only four cases that systematically use such manuals, and five in which this is sometimes done. In addition, only one organisation reported to frequently conduct meta-evaluations. In terms of explanations, the presence of an evaluation unit proved to make a strong difference to have quality measures systematically embedded. Other contributing factors are the provision of evaluation trainings, demand for evaluation from politicians, and organisational stability.

### **3 Societal dissemination/acceptance (social system)**

#### **3.1 Institutionalised use of evaluations by civil society**

While evaluation culture and practice are gradually maturing in Belgium, evaluation does not yet take a central place in decision-making. The proportional election system and the coalition governments resulting out of this, imply that coalition agreements are the ultimate reference for decision-making. The consensus-style of policy making also entails a lot of political bargaining between parties. Adding to this the strong partitocratic culture in Belgium, it is no surprise that there is only little room left for an active use of referenda or decision-making on a communal basis. The cases where evaluation can underpin these kinds of processes are rare by default. An exceptional example has been the referendum on a scenario to solve the mobility gridlock around the city of Antwerp. For years, this highly salient and longstanding issue was characterised by a closed decision-making attitude by the regional government. Gradually however, action groups came into the policy arena demanding for other alternatives and calling to approach the issue from a broader perspective (quality of air, liveability, ...). To substantiate their arguments, the action groups made active use of studies and ex ante evaluations of possible tracks for the ring road. They even managed to collect resources to commission new ex ante evaluations. Only recently an agreement between the government and the action groups was reached, which was applauded as a first step toward a sustainable and foremost supported solution.

This being said, civil society organisations and to a lesser extent private organisations or citizens do use available evaluation results in discussions with the public authorities, or in lobbying activities. We tend to argue that the main function in using evaluations is the enforcement of interests. Not seldom however, do different actors in the policy arena ‘read’ the same evaluation from a different perspective. Depending on their particular interests, stakeholders sometimes emphasise different parts of the evaluation findings and conclusions. We can name anecdotal examples in which a minister communicates evaluation findings and announces next steps, upon which civil society organisations react with their own recommendations. The latter being partly based on the evaluation findings, partly on the organisations’ general strategy and agenda. Election times are typically also a period in which civil society organisations resort to evaluations. In their so-called memoranda to the incoming government, they regularly refer to evaluations to underpin proposals. Furthermore, in line with Belgium’s neocorporatist legacy, civil society is represented in a system of advisory boards to the government. Although the advisory system is under pressure and reform (Fobé et al. 2017a), it remains a forum in which evaluations are used when formal advice is being drafted.

As for the evaluation process itself, citizens or civil society organisations have been involved in several evaluations during the information gathering stage. Their engagement in the planning, analysis or reporting stages is, yet, far from common practice. An exception is the field of international development, with a formal requirement to involve relevant actors throughout the evaluation process. Worth mentioning also are the EIA, where the participation of civil society is actively looked after. The early notification stage includes a mandatory consultation round, in which the predetermined relevant advisory bodies (i.e. administrative divisions related to environmental policy) are consulted. During this step also the general public gets the opportunity to participate and to formulate suggestions and remarks concerning the scope of the ex ante evaluation. An in-depth multiple case study (De Peuter et al. 2010) revealed the following context elements as being decisive to trigger public participation: the historical salience of the issue, the organisational capacity and availability of time to react. Also, the visibility of and know-how on the subject of environmental assessment turned out to be leverages, yet they seemed to be of a less essential nature for the degree of participation. Other contributing factors are the quality of the evaluation process, that is the quality of communication on participation opportunities, the quality of background documents available in the notification stage, as well as the organisation of extra initiatives to inform and consult stakeholders. Especially civil society organisations, rather than individual citizens, turn out to participate in these evaluation processes. Thematic interest groups have an operational reflex to monitor decision-making processes and have the capacity to more quickly react when impact assessment processes are launched. Some interest groups use the notification stage systematically as an opportunity to voice their concerns and suggestions. Having relevant topical expertise is a key in this regard. Individual citizens and ad hoc committees, in contrast, are often not acquainted with the complexity of the dossier; and with the procedures for participation. As a result, intermediary organisations often take up the role of spokesmen or substitute of the individual citizen in participation processes.

### **3.2 Public perception and discussion of evaluation and evaluation findings**

In most cases external evaluation reports are actively made public by the commissioning administration, on condition of ministerial approval. Evaluation reports are usually disseminated via the relevant governmental websites; or (to lesser extent, and if applicable) via the organiser of events at which evaluation findings are announced. Whether in-house evaluations are also made available is difficult to assess. Discussion of evaluation findings in media is not common practice and is usually restricted to these policy issues that are salient, or in case deficiencies in public sector functioning are at stake. Personal interest of journalists and competing news items also matter in this regard. A general discourse on the benefits of evaluation, the quality of



evaluation, or the professionalisation of the evaluation function does not really exist. These items are addressed ad hoc in evaluation working groups within government administrations, or during events organised by the evaluation networks (see below).

### **3.3 Civil societies demand evaluations**

The demand for evaluation mainly comes from the executive, herein often triggered by evaluation clauses in legislation, or following up on requests from stakeholders. In the neo-corporatist and consociationalist context of Belgium, such requests are usually voiced in the context of negotiations, advisory boards and parliamentary hearings. Public calls for evaluations are only made when these 'internal' channels are felt not to result in a satisfactory response from policy-makers. Again, requests predominantly come from organised civil society organisations; given their competitive advantage in monitoring the policy agenda and decision-making processes.

## **4 Professionalisation (system of professionalisation)**

### **4.1 Academic study courses and further training**

The professionalisation of evaluation takes different shape in the two main language communities. In both communities, the increased anchorage of evaluation in the public sector has been gradually matched by an expansion in supply and diversification in evaluation courses and training. At university level, there is no full-fledged master or PhD degree in evaluation as it exists in other countries, apart from a specialised Master of Development Evaluation and Management, offered at Antwerp University (Flanders). As its name suggests, this master focuses on the evaluation of development effectiveness, and is populated by a strong international student body. Most other university-level initiatives are organised by public administration departments, such as the inter-university certificate program in evaluation and the evaluation trainings provided at the recently established School for Public Administration; both are to be situated in Wallonia. In Flanders, courses in policy evaluation are included in public administration master programmes, such as the Master in Public management and Policy at KU Leuven, or the Master in Public Administration and Public Management at Ghent University. The KU Leuven Public Governance Institute further organises a five-day programme in evaluation for civil servants. In addition to the supply at public administration departments, evaluation related issues are covered at Faculties of Medicine (f.i. the course Evidence based Health Care of the Master of Health Care Management and Policy at KU Leuven) or at Educational Studies' Departments (f.i. KU Leuven organises an

annual workshop in Systematic Review). The evaluation discourse and the associated trainings in these disciplines develop along separate trajectories, though. There is hardly any exchange between them. In terms of professional training, also the evaluation associations play a major role in Belgium (see below). Furthermore, initiatives for the exchange of evaluation related information have gained ground within the civil service itself. These are often more of an ad hoc nature and are usually restricted to the civil servants of a specific policy domain only.

## 4.2 Profession/ discipline

In Belgium, there are two evaluation associations that target the different language communities. The French speaking *Société Wallonne de l'Evaluation et de la Prospective* (SWEP)<sup>2</sup> has been established in 2000 and aims at promoting evaluation and foresight in Wallonia to improve decision- and policy-making. It does so through conferences, seminars and trainings (Fobé et al. 2017b). A main impetus leading to the establishment of SWEP were the compulsory evaluations that had to be conducted in the framework of the EU structural funds programme. In Flanders, the Flemish Evaluation Association (*Vlaams Evaluatieplatform*, VEP) exists since 2007. Convinced of a need for more evaluation exchange, several academics mobilised other evaluation stakeholders including civil servants from different government administrations and the Court of Audit and policy advisers from civil society organisations. The launch of VEP anticipated an increased demand of evaluation following the attention for evaluation in the above-mentioned administrative reform framework, launched in 2003. VEP and SWEP provide a forum for all kinds of evaluation related stakeholders: civil servants, policy advisors, academics and actors from the private and non-profit sector. To be clear, the external evaluation market in Belgium is mostly dominated by scientific research institutes and consultancy firms. There are only a limited number of freelancers active in evaluation. Whereas federal civil servants are welcome to attend the activities of VEP and SWEP and both associations sometimes cooperate in the registration for events, they participate to much lesser extent. Most communication with members takes place via social media or email lists. The associations do not have their own journal, although VEP has close connections with the Flemish Journal of Public Management (*Vlaams Tijdschrift voor Overheidsmanagement*), as both initiatives are part of the Flemish Association of Management and Policy (*Vlaamse Vereniging voor Bestuur en Beleid*). Furthermore, several board members represent VEP in the editorial committee of the Dutch e-journal *Beleidsonderzoek Online* ('Policy Research Online'). The networks have played a major role in capacity building of civil servants

---

<sup>2</sup> In February 2018, SWEP announced a withdrawal as active platform mainly due to organisational capacity issues. It remains to be seen whether this constitutes a fundamental turn for network practice in Wallonia.

involved in evaluation, especially for the Walloon and Flemish public sector. There are no specific Belgian standards for evaluators. Neither is there a certification system for evaluators; or a specific evaluation professorship.

### **4.3 Compliance to standards and quality obligations**

The evaluation associations in Belgium have not developed specific standards or guiding principles but rely on existing standards in other countries (e.g. the Swiss, American, etc.). In absence of a certification system for evaluators, it is up to the members to decide whether to follow such standards. We do not exclude the possible existence of standards of guiding principles within certain disciplinary professional associations, but these are not actively communicated by the main evaluation associations. Of course, clients do request a certain level of evaluation quality for commissioned evaluations, but they usually not apply certain quality grids to evaluate proposals. Also, evaluators—at least to our knowledge—do only seldom rely on specific standards, nor do they systematically invest in meta-analyses of their evaluations. There is still a way to go before full maturity is reached, in this respect.

## **5 Conclusion and outlook**

Belgium has often been labelled as a latecomer compared to other OECD countries, demonstrating increased attention to evaluation only in the (late) nineties. In this chapter we focussed on the question how the country has evolved in terms of evaluation institutionalisation, and where it stands anno 2017.

Despite being a federal country in which internal and external impulses for policy evaluation vary between levels of government and also between its regions, the overall balance is that major progress in evaluation institutionalisation is observable. Generally speaking, evaluation culture has clearly matured in the last 2 decades, with the establishment of two evaluation associations; large scale public sector reforms directly or indirectly anchoring evaluation in the policy cycle; an increasing debate on evaluation within and by parliament; a growing number of references to evaluation in policy documents and coalition agreements; more evaluation activity; the Court of Audit shifting part of its audits to evaluation of policy results; and an extending supply of training in evaluation. Yet, as the international peloton also seems to keep the pace of maturing (Jacob, Speer and Furubo 2015) Belgium probably has not compensated for its slow start.

One should be careful, however, in sketching a general country-wide picture. It should be clear, for instance, that the level of institutionalisation both in the executive and legislative branch is less pronounced at the federal level compared to the regions. Furthermore, within the administration, institutionalisation dynamics (and hence differences) are visible mainly along sectoral or organisational paths. Indeed, with the exception of

the Walloon region, evaluation capacity remains highly decentralised and scattered across and within policy domains. With several experts, we agree that the islands and silos, along which evaluation capacity develops, endanger the future development for government wide investments in evaluation capacity and incorporate risks as to its sustainability. The country specific partitocracy adds to this risk and can foster a too narrow association of evaluation with accountability. On the other hand, intensifying dynamics for coproduction of policy and interactive policy-making, for instance, and a further decrease of neocorporatist legacies, may push the pendulum back towards evaluation for rationalistic and learning-based purposes. Time will tell which trends will prevail in the longer run.

## References

- Brans, M. & Aubin, D. (Eds.). (2017). *Policy analysis in Belgium*. Bristol: Policy Press.
- Brans, M., Pelgrims, C. & Hoet, D. (2005). Politico-administrative relations under coalition governments: the case of Belgium. In B. G. Peters, T. Verheijen & L. Vass (Eds.), *Coalitions of the unwilling? Politicians and civil servants in coalition governments* (pp. 207–235). Bratislava: NISPAcee. The Network of Institutes and Schools of Public Administration in Central and Eastern Europe.
- De Maesschalck, F., Holvoet, N. & Hooghe, I. (2016). Beleid, gebruik en invloed van evaluaties in de Belgische ontwikkelingssamenwerking: een blik op de Dienst Bijzondere Evaluatie. *Vlaams Tijdschrift Voor Overheidsmanagement*, 9(2), 55–62.
- De Peuter, B. & Pattyn, V. (2016). Waarom evalueren beleidsmakers? Een longitudinale analyse van motieven voor beleidsevaluatie in Vlaamse ministeriële beleidsnota's. *Bestuurskunde*, 25(2), 32–45.
- De Peuter, B., Houthaave R. & Van Der Linden M. (2010). *Case studies mer-processen bij plannings en vergunningstrajecten*. Leuven: Instituut voor de Overheid.
- Deschouwer, K. (2012). *The politics of Belgium: Governing a divided society*. Basingstoke: Palgrave Macmillan.
- Desmedt, E., Pattyn, V. & Van Humbeeck, P. (2016). Beleidsevaluatie vandaag : Een voorzichtige balans. *Vlaams Tijdschrift Voor Overheidsmanagement*, 21(2), 63–69.
- Dewachter, W. (2001). *De mythe van de parlementaire democratie: een Belgische analyse*. Leuven: Acco.
- Dienst Administratieve Vereenvoudiging (2014). *Regelgevingsimpactanalyse. Inleiding tot praktische oefeningen*. Brussels: DAV.
- Fobé, E., Biard, B. & Schiffino, N. (2017a). Policy advisory bodies in Belgium. In M. Brans & D. Aubin (Eds.), *Policy Analysis in Belgium* (pp. 151–170). Bristol: Policy Press.

- Fobé, E., De Peuter, B., Petit Jean, M. & Pattyn, V. (2017b). Analytical techniques in Belgian policy analysis. In M. Brans & D. Aubin (Eds.), *Policy analysis in Belgium* (pp. 35–56). Bristol: Policy Press.
- Furubo, J., Rist, R. & Sandahl, R. (2002). *International atlas of evaluation*. New Jersey: Transaction Publishers.
- Jacob, S. (2004). L'institutionnalisation de l'évaluation des politiques publiques en Belgique: entre balbutiements et incantations. *Res Publica*, 4, 512–534.
- Jacob, S. (2005). *Institutionnaliser l'évaluation des politiques publiques: étude comparée des dispositifs institutionnels en Belgique, en France, en Suisse et aux Pays-Bas*. Brussels: Peter Lang.
- Jacob, S., Speer, S. & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21(1), 6–31.
- OECD. (2015). *Impact assessment in Belgium, federal government*. Paris: OECD.
- Pattyn, V. (2014). *Policy evaluation (in) activity unravelled. A configurational analysis of the incidence, number, locus and quality of policy evaluations in the Flemish public sector*. KU Leuven: Public Management Institute.
- Pattyn, V. (2015). Explaining variance in policy evaluation regularity. The case of the Flemish public sector. *Public Management Review*, 17, 1475–1495.
- Pattyn, V., & Brans, M. (2013). Outsource versus in-house? An identification of organizational conditions influencing the choice for internal or external evaluators. *Canadian Journal of Program Evaluation*, 28(2), 43–63.
- Pattyn, V., & Brans, M. (2014). Explaining organisational variety in evaluation quality assurance. Which conditions matter? *International Journal of Public Administration*, 37(6), 363–375.
- Pattyn, V., & Brans, M. (2015). Belgium. In D. A. Bearfield, E. M. Berman, & M. J. Dubnick (Eds.), *Encyclopedia of public administration and public policy* (3rd ed., pp. 1–4). Boca Raton, Florida: CRC Press.
- Sociaal-Economische Raad van Vlaanderen. (2015). *Advies: Tien denksproten voor ex post decreetsevaluatie in en door het Vlaams Parlement*. Brussel: SERV.
- Speer, S., Pattyn, V. & De Peuter, B. (2015). The growing role of evaluation in parliaments: holding governments accountable? *International Review of Administrative Sciences*, 81(1), 37–57.
- Varone, F., Jacob, S. & De Winter, L. (2005). Polity, politics and policy evaluation in Belgium. *Evaluation*, 11(3), 253–273.

Verlet, D., Lemaître, J., & Carton, A. (2015). *Beleidsvaluatie binnen de Vlaamse overheid. Een Overzicht van de resultaten uit de bevraging van de leidinggevenden*. Belgium: Presentatie Studiedag Vlaams Evaluatieplatform. <http://www.evaluatieplatform.be/VEP/programma/20151217.htm>. Accessed on 29 January 2019.