Toward a mature evaluation culture

in Belgium and Flanders?

A search of drivers, recent developments and challenges.

Paper for the *Symposium "Policy and programme evaluation in Europe: cultures and prospects. Strasbourg, July 3 & 4, 2008.*

Bart De Peuter

Public Management Institute

University of Leuven

Belgium

Contact:

Bart De Peuter
bart.depeuter@soc.kuleuven.be
Voice: +32 16 32 32 73

INSTITUUT VOOR DE OVERHEID

Public Management Institute
Faculty of Social Sciences
University of Leuven

Parkstraat 45, po box 3609
3000 Leuven
Belgium

io@soc.kuleuven.be
Voice: +32 16 32 31 80
Fax: +32 16 32 32 67
www.publicmanagement.be

1. Introduction

Some authors (e.g. Furubo, Rist and Sandahl, 2002) have developed indicators to assess the existence and maturity of an evaluation culture in a particular country or region. The indicator set of Furubo a.o. comprises in fact different measures of institutionalised practice.In this paper we want to take stock of the evaluation culture in Belgium and in one of its regions, Flanders in particular.

First we examine the drivers for governments to evolve toward an evaluation practice (section 2). In a next section (3) the Belgian case will be adressed in this regard. Also, we will apply the indicators of Furubo et al. to explore the maturity of the evaluation culture in Belgium and Flanders. In a concluding section (4), we advance some challenges. They are meant as input for further debate and thought on future directions in which evaluation culture and practices can evolve.

2. Drivers toward an evaluation culture

In this section we outline some key drivers with regard to the increased attention for evidence-based policy in general and evaluation in particular. We discern between two types of drivers: on the one hand general factors and trends which shape the knowledge society and on the other hand drivers that are characteristic for a public sector context.

2.1 General trends towards a knowledge society

Evidence-based policy can be seen as a public sector translation of a set of more wider trends in society which has generally evolved towards a knowledge society. A first fundamental driver is the technological evolution during the past two decades. The world wide web and wireless communication technology caused a major change in how people and organisations can interact. A second fundamental driver of the knowledge society is the spread of democratic values throughout the world and within societies. Issues of inequalities, equal chances and diversity appear on top of many political agenda's. These two basic trends, technological innovation and democratisation- led toward a world perceived as a 'global village'.

Consequently, the current society can be considered as a knowledge society; a broad range of social issues and phenomena as well as experience and organisations became much more 'knowledge-intensive'. The aspect of knowledge as such is not new however, but the societal role of knowledge has changed dramatically. A wider spread of education and rapid access to channels of information diffusion among much more actors creates more room for debate and contestable arguments.

The impact on governance and policy-making is that they became more and more knowledge-intensive. Together with a tendency of growing complexity and a rapid pace of change, the response time of governments and their governmental capability are put to the test. Knowledge production and management within government become increasingly important points of interest, while what constitutes knowledge has been subject of debate as well, between advocates of respectively a quantitative or qualitative approach which relate to different paradigms of positivism and constructivism (Worthen, Sanders and Fitzpatrick, 2003).

2.2 Public sector specific drivers

In addition to the general trends described above, a number of factors which are more specific to the public sector are drivers towards an evidence-based policy.

An initial driver has been the *fiscal stress* governments have to deal with (Bouckaert & Auwers, 1999). This often resulted in an attempt to decrease budget deficits. When there is no information available on the relative importance of some expense categories, budget cuts often result in linear saving operations according to an arbitrarily fixed percentage. Such an input approach does not take into account social needs. With the concept of value-for-money the approach changed and the focus of measurement information shifted towards the quantity and quality of outputs and outcome.

A second driver is the set of institutional reforms within government which lead in most countries toward more *decentralisation of policy implementation*, by the creation or transformation of autonomised agencies for specific policy issues. The proliferation of entities and organisations increased not only the number of policy initiatives, which entails a risk of decreasing transparency of management and policy. Policy cycles become decoupled: core departments take up responsibilities of policy planning and preparation whereas agencies with managerial autonomy are entrusted with implementation. The following figure, taken from Bouckaert and Verhoest (2003), displays 3 positions for OECD countries. Position 1 reflects the 1980s when monolithic ministries tried to include all aspects of a policy cycle and remained in quadrant I. Between 1980 and 1995 there is a combined evolution of organisational proliferation based on the creation of autonomous or quasi-autonomous agencies, devolution, and decentralisation on the one hand, and separating stages of the policy cycle on the other hand. This is reflected in a shift to position 2. However, more recently in many countries it was realised that specialisation and autonomy could only yield value added if there was sufficient coordination (position 3). As a result, reconsolidating mechanisms where triggered varying form hierarchy type mechanisms (HTM), market type mechanisms (MTM) to network type mechanisms (NTM).

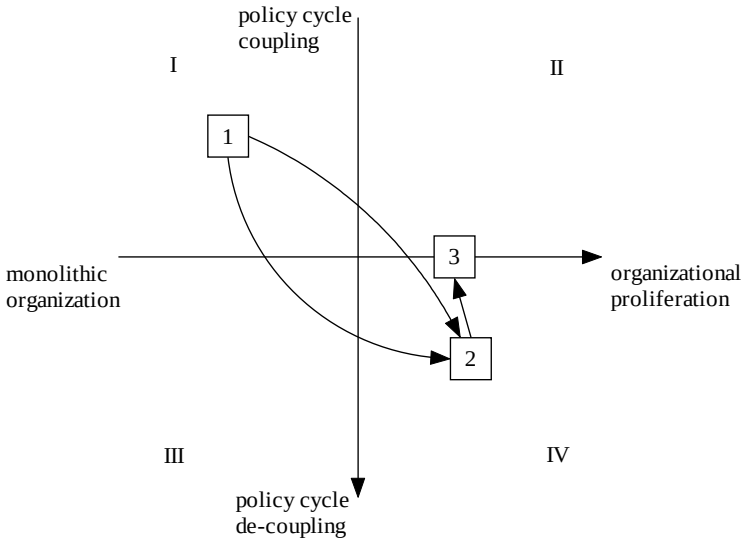

Figure 1: Specialisation of organisations and policy cycles: trends in OECD countries

Obviously, these shifts have also an impact on the feedback lines of information within the governmental machinery. The success of coordination and control mechanisms is largely dependent on the way the supply and demand of policy information are matched for these steering and accountability relations within and between public sector organisations.

Not only within the same governmental level there is horizontal and vertical institutional division. A third driver has to do with the trend toward more *intergovernmental policy-making*. Some studies of central-local relations take a normative-legalistic approach and tend to draw

a sharp difference between the local and central sphere of government based on a dual polity-model (Wayenberg, 2004). However, there seems to be quite some evidence that this proposition is not longer accurate to describe the reality in this regard. Rhodes (1999) for example pointed at a general trend in Western-Europe toward intergovernmental policy-making: policy issues are more and more handled in concerted action by central and local government. In fact, he referred to the existence of a two-way traffic or interaction between central and local government, describing it as central-local networks, being 'service-specific professional-bureaucratic complexes'. Several authors, for example Batley and Stoker (1991), developed a typology of network models based on different central-local dependency relations, stating that the different types of models can be present within the same country.

In recent studies several authors concluded more specifically on a general trend toward partnership as the prevalent pattern of central-local relations in Europe (Baldersheim, 2002; Banner, 2002, John, 2000; Wayenberg, 2004). Country-differences are still eminent, but relate to different starting conditions between countries in the light of reform trajectories (Pollit and Bouckaert, 2004). Regardless of country(group)-specific intergovernmental patterns, central government always fulfils a twofold role in its relations with local government. On the one hand, central government plays a task delineating, regulating and delegating role vis-à-vis the local tier of government consisting of territorial decentralised governments. On the other hand, central government plays an enabling and supportive role vis-à-vis the local tier of government, consisting of authorities with local self-government (De Peuter & Bouckaert, 2004).

Not only central and local levels of government witness more intergovernmental policy-making. Also central governments and the European level are becoming more and more intertwined. Policy-making and implementation in the EU context is a continuous interplay between the European institutions (Council, Commission and Parliament), the member state governments, regional actors and interest groups. Over the years the European level expanded its role in a number of policy areas and became involved in new policy areas, introducing again new alternative forms and instruments of intergovernmental policy-making in addition to the "community method", e.g. by The Maastricht Treaty of 1992. The creation of diverse and alternative methods of European governance across the wide range of policy areas stems from the fundamental mechanism underpinning the integration process. This mechanism can be visualized by a pendulum swinging between two magnetic fields, the one country-based and the other trans-national (Wallace & Wallace, 1999). The pendulum's moves depend on the varying strengths of these two fields and are different according to the policy area.

Thus, different layers of government have become increasingly intertwined in the policy making process. However, the division of tasks between the levels of government may differ across policy fields and will rarely seamlessly correspond with a clear distinction between (upper level) policy formulation and (lower level) policy implementation. In fact, in a *multi-level governance context* different levels also share responsibilities throughout the classical scheme of the policy cycle (preparation – decision – implementation – evaluation).

A fourth driver is the increased *complexity of policy-making*. Policy issues become more and more tied up with one another, involving many actors which have to cooperate in dealing with complex societal problems across geographic borders and administrative boundaries. This horizontal interlacement of policy fields results in a need for a broader view on the policy context and thus, insights in the complex environment in which policy problems arise and have to be solved.

2.3 Reform context of professionalisation and inter-activeness

When considering the drivers of change behind the knowledge society, public sector organisations are confronted with a reform context of professionalizing and inter-activeness when developing an adequate response to the challenges the knowledge society puts forward. On the one hand public sector aspires to modernise and adapt its policy-making capacity. Professionalizing structures, procedures, personnel and values occurs with the overall intention to develop a better, i.e. effective policy to tackle the identified social problems. On the other hand, there is an increasing demand of more interactive policy-making processes, with consultation and participation of external actors, and/or individual citizens (Vancoppenolle & Brans, 2004). In this two-fold reform context, the role of policy information to build evidence for policy decisions is quintessential.

Evaluation, together with monitoring and foresight, constitutes an important lever to underpin policy decisions throughout the policy-cycle. Given the drivers stated above, public policy-makers are confronted with internal and external pressures to take decisions that are well informed. We can discern between two general rationales from evaluation are then deemed important: to strive for a better policy management and for reaching consensus between different stakeholders. Under these perspectives a whole range of more concrete purposes can be put. When we look at the broad range of motives stated in literature and practice, we can regroup them. (See e.g. Chelimsky & Shadish, 1997). De Peuter, De Smedt and Bouckaert (2007) highlight a set of four recurring purposes: (1) strategic planning support, (2) implementation improvement, (3) accountability and (4) policy learning. Taken together, these evaluation purposes to build an evidence base to support policy cover in fact the complete policy cycle. It means that throughout the policy cycle evidence can be gathered for different purposes.

3 Evaluation culture in Belgium and Flanders: drivers, trends and challenges

3.1 Drivers

Belgium and Flanders have introduced evaluation thinking and practice mainly under external pressure. In fact, Belgium positions itself as a latecomer. This needs some clarification. In Europe evaluation practice and culture have spread slowly and in two waves.

Only a limited number of countries constitute the first wave, at the end of the 1960s, beginning of 1970s. A mix of elements created for them an *internal* pressure which was at the time not present in other countries. Furubo (2002) states a greater political-cultural openness towards evaluation, from a political tradition of strong belief in the ability of government to steer society, favouring rationalistic attitudes. Further, civil servants were familiar with social scientific thinking and had intensive contacts with the US where evaluation had already more developed. Also, in the early adopting countries, there existed a perceived need for evaluation, which correlated with the role of the public sector. Indeed, the early adopters did have a higher level of public expenditure at that time compared to countries who belong to a second wave of developing an evaluation praxis.

The second wave came in the 1990s. Interestingly, second wave countries faced rather *external* pressure to introduce evaluation in the public sector. At the end of a second wave with which evaluation has been spread among European countries.

The same goes for Belgium, where the momentum for increased attention to policy evaluation only came when the second wave was already slowing down. For Belgium, four drivers can be discerned.

A first impetus in Belgium came with the spread of the *New Public Management ideas*. Especially the accountability purpose to evaluate came to the front. Contrary to the first wave countries, in Belgium evaluation thinking was no strategic choice but was introduced as part of a broader package of reform and modernisation concepts.

A second important source of external pressure came from the *European Union*. Many European policy programmes and initiatives require an active participation of different governmental levels in the member states. The European Union has developed a strong evaluation function and culture within the Commission, which further radiates to the national and regional governments. The European Commission, frequently in the role of policy initiator and financer, is strongly interested in the efficient use of resources and the level of goal achievement. This mechanism requires member states to participate in an evaluation praxis. In Belgium, numerous federal and regional administrations, e.g. in the environmental, agriculture and fisheries sector, labour policy and regional development, have practical experience with 'evaluating for Europe' (De Peuter & De Smedt, 2006).

A third driver is related to the numerous *reform processes* which currently are undertaken within the different governmental levels in Belgium. Varone and Jakob (2003) pointed at the importance of additional efforts for the institutional embedding and development of evaluation. Not only it is necessary to undertake quality and solid evaluation processes. Also institutionalisation increases the political and social visibility of evaluation results and of the related policy improvements.

The attention given to evaluation in the framework of several reforms points at a fourth driver: the growing belief *among policy-makers* that evaluation can deliver added value to policy. In this perspective, external pressures can trigger internal incentives which accumulate toward a more mature evaluation culture.

3.2 The maturity of evaluation culture in Belgium and Flanders

It is difficult to measure how strong the existing evaluation culture in a given country or region is. Furubo (2002) has developed a set of nine indicators to assess whether and to what extent an evaluation culture has reached full maturity. In this section we will use these indicator set as a guidance for an empirical snapshot of recent trends in this case study.

A first indicator in this regard is whether *evaluation takes place in many policy domains*. This is true for Belgium and Flanders. Although evaluation may be relatively limited still, activities are not limited to one or two policy domains. Empirical evidence for this we found in a number of sources, i.e. policy (planning) documents relating to several policy sectors referring to existing or planned evaluation procedures and activities, announcements of public tenders and evaluation reports published on websites of governmental organisations. Of course, we can discern leaders and laggards as far as evaluation practice is concerned. For example, in Flanders, policy sectors such as education, labour and environment have a longer tradition and broader experience with evaluation than other sectors.

A second indicator constitutes the *supply of evaluators* specializing in different disciplines who have mastered different evaluation methods and who conduct evaluations. This condition does not yet fully apply to Belgian and Flemish practice. On the one hand, the number of evaluation functions and evaluators within government organisations is limited but increasing. Further more, the evaluation procedures embedded in legislation (see further) tend toward convergence of methods. On the other hand, the supply of external evaluators is partly covered by academic research centres which master a broader range of evaluation methods and have persons with varying educational backgrounds, such as political and

social sciences, law, economics, etc. Still, the number of private sector consultant organisations offering evaluation services remains limited. There is anecdotic evidence that some consultants are often the sole applicants for an evaluation put out to tender by the Flemish government.

A third indicator Furubo mentions with regard to an evaluation culture, is the existence of a *national discourse concerning evaluation* in which more general discussions are adjusted to the specific national environment. The alternative is that issues and questions such as utilisation, organisational structures, systems for training evaluators, and so on do not result from a country's or region's own experience and discourse, but is totally based on 'imported goods'. At present, one cannot speak of a national discourse on evaluation, since there is no comprehensive discourse addressing regularly all aspects of the evaluation process. As far as methods are concerned, these are mostly based on imported ideas or procedures. This can be explained by the external pressure to evaluate stemming from the European Union. On the other hand, for specific discussions reference is inevitably made to domestic contexts with regard to organisational structures.

Fourthly, an evaluation culture is in place when there is a *profession with its own societies* or frequent attendance at meetings of international societies and at least some discussion concerning the norms or ethics of the profession. There is no Belgian evaluation association or society. However, on the regional level, initiatives have been taken or are in a planning stage. In Wallonia, the 'Société Wallonne d'Evaluation et de Prospective' (SWEP) was established in 2000 with the following mission statement:

- To organise a methodological and ethical reflection oriented to the actors directly involved in evaluation and prospecting activities;
- To improve the diffusion of evaluation and prospecting culture among citizens and policy-makers;
- To encourage policy-makers to learn and to support evaluation and prospective activities in accordance with acknowledged ethical values and technical requirements.

In Flanders, a Flemish Evaluation Platform ('Vlaams Evaluatieplatform' – VEP[1]) has been established in December 2007. This network, like other evaluation associations and societies, is meant to be a place where knowledge, experience and expertise can be shared among persons who professionally are involved in an evaluation process. The target group comprises civil servants from all governmental layers, academics and actors from the private and not for profit sectors. It is prepared by a steering committee composed of researchers from universities, civil servants from different policy domains and the Court of Audit of Belgium. The objectives of the Flemish Evaluation Platform are:

- to build, strengthen and diffuse an evaluation culture;
- to enhance the evaluation capacity of the actors involved;
- to increase quality of policy evaluation;
- to stimulate the influence and use of policy evaluation.

A fifth indicator concerns *institutional arrangements in the government* for conducting evaluations and disseminating their results to decision makers. This criterion attempts to take

[1] For more information, see its website: www.evaluatieplatform.be

into consideration permanent arrangements or systems whereby evaluation initiatives are commissioned to different evaluators and at the same time arrangements are developed for ensuring that the evaluations conducted are put to suitable use. This is a form of guarantee that utilization – at least in formal terms- will take place.

A common trend for the different governmental levels in Belgium is indeed to embed the requirement to evaluate into the policy cycle by legislation. On the federal level the so-called sustainable development effect assessment was introduced in 2003. The test applies to all important legislation and is a next step within this transversal policy field after the creation of entities within each federal department. The introduction of a similar test is currently prepared looking at consequences of measures with regard to gender equality.

Also on the regional level a number of evaluative 'tests' has been introduced. They are in fact ex ante evaluations of a scheduled legislation on topical points of attention, e.g. the child effect report, the poverty test, and the compensation measure to limit administrative burdens for citizens and enterprises. The latter principle states that when new administrative burdens are created by new regulation, this must be compensated by abolishing other administrative burdens.

Some evaluation requirements have been inserted into national or regional law to implement European directives. The strategic environmental assessment (SEA) for projects and programmes is an example. In 2004 the regulation impact assessment procedure (RIA) was introduced providing a framework for some of the tests mentioned. RIA is a structured analysis of the envisaged objectives and expected positive and negative effects of a planned regulation, while compared to alternative approaches such as self-regulation or covenants.

In Flanders, evaluation has been generally anchored in the policy cycle by the framework decree introduced in the light of the comprehensive reform of the Flemish administration, called 'Better Administrative Policy' (Beter Bestuurlijk Beleid). According to the Framework Decree (Vlaams Parlement, 2003) the departments are responsible for policy supporting tasks. The explanatory memorandum of the decree states among other things the evaluation of policy implementation and, more in particular, the effectiveness of the instruments used, the relations between output and outcome. The autonomized agencies have to generate input by means of relevant policy and managerial information for policy evaluation. The order in pursuance of the decree states the tasks of departments more in detail. They must support the minister in:

"1. the elaboration of his/her policy, which means the department has a task of policy preparation and evaluation;

2. the steering and follow-up of policy implementation as carried out by the agencies."

Regarding policy preparation and evaluation for the respective Ministers, one of the tasks for departments is to provide permanent monitoring and environmental scanning of the policy domain (included the management of data bases and the analysis of statistical information). Secondly they are entrusted with the evaluation on macro level of the policy implementation (instruments used, outcomes, etc.) with a view to an eventual adjustment of the policy or the steering of agencies. The management of policy relevant scientific research and the utilization of research results is another task. Departments are also responsible for the development of tools for policy planning (models, scenario analysis, benchmarking,…). Finally, departments must elaborate proposals with regard to policy formulation (regarding instruments, resources needed, financing mechanisms, periodically needed policy and management information, reporting requirements, accountability and control mechanisms) and concerning the review of policy proposals.

The decree on the financial accountability states the policy effect reporting (on the level of a policy domain as a whole): a report which examines and evaluates the achievement of policy options during the past financial year by means of indicators and key figures which are linked to strategic policy objectives (social outcomes) and operational objectives (policy output) and which are stated in the policy document.

A screening of policy documents of the Flemish government for the period 2004-2009 examined to what extent and how the principles of the reform 'better administrative policy' were translated into these plans (Conings e.a, 2005). One of the conclusions is that there is a general need for a more evidence-based policy. Most of the documents also contain different references to scheduled evaluations, the development of evaluation tools and the evolution toward evidence-based policy. The analysis also took stock of the number of indicators mentioned with regard to public policy and management; on average 5 input indicators, 5 output indicators 10 outcome indicators and 11 environment indicators were stated in each policy document (n=30).

Also as far as the local governmental level is concerned, we find a reinforced attention paid to strategic planning together with reference to policy evaluation in the newly introduced Municipality Decree and Province Decree (2005). The municipal secretary and the clerk of the province received responsibilities on the field of evaluation of policies, and a management team has to support the coordination of services during the evaluation processes. At this moment it is too early to assess the impact in practice of these legislative incentives.

Nevertheless, the embedding in legislation of evaluation related tasks and reporting are an evident sign that evaluation becomes more institutionalised. Another indicator lies within the organisational reforms within government administrations. On all levels the institutionalisation of monitoring and evaluation is also visible by the creation of specialised cells or task forces. They are important levers for an internal capacity aimed at the planning, application of follow-up of processes and tools to build an evidence base to inform policy formulation.

In the context of inter-governmental relations, policy information is shared in the framework of several covenants between the central and local government. Covenants are mostly linked to subsidies to local governments which they receive in exchange of planning and reporting efforts. The data which are in this way provided to central government constitute a potential important source of policy information.

The Belfort principle ensures that, whenever central government enacts legislation, the consequences for local governments in terms of personnel, operational budget and investments are determined on beforehand.

A sixth indicator Furubo introduced is the presence of *institutional arrangements in Parliament* for conducting evaluations and disseminating them to decision makers. When applied to Belgium we observe no major initiatives in this regard. This does not mean members of parliament are not interested in evaluation. Frequently they formulate a need to evaluate policies and, on the regional level, the results of the regulatory impact assessments are discussed. But there are no institutional arrangements by which evaluation is anchored parallel to those within government administration. It has been demonstrated that there is interest from members of parliament with regard to quantitative policy information (Van Dooren (2003).

Another indicator to assess the maturity of an evaluation culture is *pluralism*, i.e. within each policy domain there are different people or agencies commissioning and performing evaluations. While this is the case to a large extent, it must be stressed that the (external) supply side on the market remains rather small.

An additional indicator stated by Furubo is the presence of *evaluation activities within the Supreme Audit Institution*. Indeed, Pollit (1999) defined four audit roles: judge, public accountant, management consultant and researcher. In the latter role the auditor creates new knowledge about the outputs and outcomes of public programs and could undertake evaluations. The Court of Audit of Belgium has seen an extension of its duty by the law of 1998 so that it became competent for a check of the good spending of public money and to verify that the principles of economy, efficiency and effectiveness are complied with.

Thus, part of examinations of the Court of Audit will also look at effects. However, while these so-called 'performance audits' take only a small share, they mostly use an indirect approach to the assessment of effectiveness, by a judgement whether the conditions for effectiveness are fulfilled, the performance and the extent to which the target group is reached (Put, 2000). Put also points to the hostile environment in which audit institutions fulfil their function, and of which they are dependent in some perspective. This makes it more difficult for them to undertake real evaluations of policy initiatives. Pollitt referred to the fact that the different roles of an auditor may be complementary as well as competing.

A last indicator refers to the focus of evaluations and the requirement that part of the evaluations carried out are not just focused on the relation between inputs/outputs or technical production. *Some public sector evaluations must have program or policy outcomes as their object* and raise such questions as whether the public interventions actually had an impact on the problems they were intended to solve. This 'condition' is fulfilled in Belgium and Flanders. In fact, the more recent initiatives to embed evaluation into the policy and legislative process described above, have a focus on effects or outcome of policy measures and programmes.

When we reconsider the set of indicators that Furubo suggests to assess the extent of an evaluation culture and its application to Belgium and Flanders, we can conclude that both on federal and regional level there is evidence of a maturing evaluation culture. Nevertheless, the practice and culture today is still immature. This is especially the case within the local governmental level.

In order to measure the pace with which an evaluation culture – and more broader: an evidence-based policy practice- is spreading, one could use time series on some of the empirical criteria we have dealt with; the number of evaluations commissioned, the number of indicators mentioned in parliamentary proceedings, policy documents and annual reports of governmental organisations.

4 Challenges

In the previous section we attempted to outline the maturity of evaluation culture in Belgium and Flanders. We conclude this paper with a number of challenges in the light of further developments in this regard.

Although Flanders has made considerable efforts in developing monitoring systems, there is still a need for more, additional and new information systems in order to support the objective of an evidence-based policy (Vlaamse Regering, 2004). This challenge was also

highlighted by a survey among ministerial advisors and top level civil servants in Flanders (Vancoppenolle & Brans, 2005). An accompanying challenge is to *match the information supply from monitoring systems with the most important evaluation activities and needs*. Planning an ex post evaluation of the effectiveness of a given policy should be prepared before the policy is actually implemented, by examining the needs of information gathering and analysis in order to be able to answer the evaluation questions raised later on.

Data exchange between public organisations and governmental levels is a crucial point of interest when we keep in mind the drivers of evidence based policy we stated earlier. Policy is less and less a matter of one organisation or one governmental level in Belgium. Although there is much desire to move to a division into homogenous competences for each governmental level in Belgium, cooperative federalism and intergovernmental policy is inevitable and this has consequences on the field of information. Intergovernmental policy requires to a certain extent also intergovernmental evaluation. This kind of evaluations entails particular challenges in each step of the evaluation process (De Peuter & De Smedt, 2006).

Another challenge yet to be dealt with is to *consolidate existing effectiveness research from different sources*. This constitutes an important dimension of a government-wide evidence based policy practice. The crucial question 'What works?' can only be adequately answered when knowledge is built up along a well planned strategy to take stock of results from several evaluation reports and scientific research.

In Flanders, some policy documents translate the demand for reports on outcome on the level of a policy domain into the development of *evaluation tools which assess goal attainment of objectives across policy domains*. Although the cross-sectoral dimension is not anchored in legislation by the overall reform 'Better Administrative Policy', this need is formulated and needs to be addressed. The challenge is here not only to find apt methods, but also to coordinate different tools and reporting instruments.

It is clear that institutionalisation of evaluation in Belgium has developed with a bias to the executive branch of government. However, in order to come to a mature evaluation culture, it will be necessary to *involve members of parliament* in a more structured way in major evaluation procedures. This is a lever to enhance evaluation use and influence.

Another area of tension is the dilemma between e*xpertise-based vs. experience-based evidence* in evaluation. The growing complexity of policy issues, problems and solutions, brings along so-called intractable policy controversies. They are characterised by fuzzy countours and boundaries, and uncertainty over which disciplines and specialisms should be taken recourse to for generating solutions. At the same time, there is a fragmentation of expert advice. Because knowledge has become pluralist, each expert can be confronted by a counterexpert. Knowledge has hence become an integral component of political discussion and argumentation.

Besides these inherent problems with professional expertise, the latter is increasingly contested by other types of expertise, particularly by the experience-based expertise of groups of citizens. When the latter is ignored, citizens' disenchantment may engender ongoing policy controversies, in which court proceedings interfere with the government's search for effective policy solutions (see for instance NIMBY cases in sectors such as the environment and infrastructure). This causes major challenges for governments to integrate different types of evidence into a workable problem definition and solution analysis. In fact, it

is a clash between the two trends we pointed at earlier: professionalisation and interactiveness.

The ultimate challenge is of course not just to have an evidence base to inform policy by evaluation, but to have an *evidence base that is used or which influences policy*. This can be strengthened by improvements of the information system itself. But to a large extent there will be a cognitive process and change needed, too. Evaluation culture and capacity influence each other. Perhaps this reflects the story of the chicken and the egg: what was (or has to be) first? But at the end of the story: we need both.[2]

[2] See in this regard another paper presented at this symposium: De Peuter Bart and Pattyn Valérie, *Evaluation capacity: enabler or exponent of evaluation culture?*

Bibliography

Baldersheim, H. (2002). 'Subsidiarity' at work: modes of multi-level governance in European countries. In: Caulfield, J. and H. Larsen (reds.) *Local government at the Millennium*, pp. 203-213. Opladen/Germany: Leske and Budrich Publishers.

Banner, G. (2002). Community governance and the new central-local relationship. *International Social Science Journal*, 54(2), pp. 217-231.

Batley, R. and G. Stoker (1991). *Local government in Europe: trends and developments.* Basingstoke: Macmillan.

Bouckaert G. & Verhoest K. (2003). *Machinery of Government: Shifting from Effectiveness of Coordination to Coordination of Effectiveness.* Paper presented at the Workshop: 'Challenges to State Policy Capacity: Global Trends and Comparative Perspectives', City University of Hong-Kong.

Bouckaert, G. and T. Auwers (1999). *Prestaties meten in de overheid*. Overheidsmanagement nr. 5. Brugge: Die Keure.

Brans M. and Vancoppenolle D. (2005). "Poilcy-Making Reforms and Civil Service Systems: An Exploration of Agendas and Consequences" In: Painter M. and Pierre J. (eds.), *Challenges to State Policy Capacity: Global Trends and Comparative Perspectives*, New York: Palgrave Macmillan, pp. 164-184.

Cabinet Office (2000). *Professional Policy-Making in the 21st Century.*

Chelimsky, E. and Shadish W.R. (1997). *Evaluation for the 21st century: a handbook* Thousand Oaks (Ca): Sage.

Chen H.-T. (2005). *Practical Program Evaluation*, Sage: Thousand Oaks.

Conings V., Sterck M., Van Dooren W. en Bouckaert G. (2005). *Beleidsnota's versus Beter Bestuurlijk Beleid. Een toetsing van de beleidsnota's aan het Comptabiliteitsdecreet en het Kaderdecreet Bestuurlijk Beleid.* Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen, 102 p.

De Peuter B. & Bouckaert G. (2004). *Een monitoringssysteem voor lokale bestuurskracht : naar de modellering van een Vlaams design.* Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

De Peuter B. & De Smedt J. (2006). *Problems of Policy Evaluation in a European Multi-Level Governance Context: The case of Active Labour Policy in Belgium*. Paper presented at the UKES / EES Joint International Conference Evaluation in Society: Critical Connections, London, UK.

De Peuter B., De Smedt J. and Bouckaert G. (2007). *Handleiding beleidsevaluatie: Deel 1: Evaluatiedesign en –management*. Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

Federaal Parlement (1998). *Wet tot wijziging van de organieke wet van 29 oktober 1846 op de inrichting van het Rekenhof*. Brussel.

Federaal Parlement (2006). *Wet tot wijziging van de wet van 4 juli 1962 betreffende de openbare statistiek*. Brussel.

Furubo J.-E., Rist R.C. and Sandahl R. (eds.) (2002). *International Atlas of Evaluation*, New Jersy: Transaction Publishers.

Jacob S. et Varone F. (2003). *Evaluer l'action publique: état des lieux et perspectives en Belgique, Série Modernisation de l'Administration*, Gent: Academia Press.

John, P. (2001). *Local Governance in Western Europe*. London: Sage.

McDavid J.C. & Hawthorn L.R.L. (2006). *Program Evaluation and Performance Measurement, an Introduction to Practice*, Thousand Oaks, CA: Sage.

Pollit C., and Bouckaert G. (2004). *Public Managment Reform*, New York: Oxford University Press.

Pollitt C. e.a. (1999). *Performance or Compliance? Performance Audit and Public Management in Five Countries.* Oxford: Oxford University Press.

Put V. (2000). "Doelmatigheidsonderzoek in het Rekenhof." In: *Vlaams Tijdschrift voor Overheidsmanagement*. 5(2-3), pp. 19-24.

Rhodes R. A. W. (1999). *Control and power in central-local relations - second edition*. Aldershot: Ashgate.

Rossi P. & Freeman H.E. (2004). Evaluation: A Systematic Approach, Seventh Edition, Newbury Park: Sage.

The Economist (9-11-1991).

Van Dooren W. (2003). *Supply and demand of policy indicators: a cross sectoral comparison.* Paper presented at the workshop' Evidence-based policy and indicator systems', London.

Van Doorn J.W.M. en Van Vught F.A. (1982). "Toekomstonderzoek: voorspellingsmethoden". In: *Handboek beleidsvoering voor de Overheid*, Alphen aan den Rijn: Samsom H.D. Tjeenk Willink, , pp.1-19.

Vancoppenolle & Brans (2004). *Ambtelijk beleidswerk in vernieuwende beleidsvormingsprocessen.* Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

Vancoppenolle & Brans (2005). *De Vlaamse beleidsvormingscapaciteit: Een (zelf-)evaluatie.* Leuven: Steunpunt Bestuurlijke Organisatie Vlaanderen.

Vlaams Parlement (09-07-2003). *Decreet Strategische Adviesraden.*

Vlaams Parlement (09-12-2005). *Decreet Provinciedecreet*.

Vlaams Parlement (15-07-2005). *Decreet Gemeentedecreet*.

Vlaams Parlement (18-07-2003). *Decreet Kaderdecreet Beter Bestuurlijk Beleid*.

Vlaamse Regering (2004). *Beleidsnota Vlaamse Statistieken, Indicatoren en Toekomstverkennende Studies voor het Regeringsbeleid*. Brussel.

Wallace H. and Wallace W. (1999). *Policy-Making in the European Union*, New York: Oxford University Press.

Wayenberg E. (2004). *Vernieuwingen in de Vlaamse centrale-lokale verhoudingen: op weg naar partnerschap? Een kwalitatieve studie van de totstandkoming en uitvoering van het sociale impulsbeleid*. Doctoral thesis, Leuven: K.U. Leuven, departement Politieke Wetenschappen.

Worthen B.R. , Sanders J.R. and Fitzpatrick J.L. (2003). *Program Evaluation: alternative approaches and practical guidelines*, Reading (Mass.): Addison-Wesley.