

# Draft Evaluation Policy Guidelines of the European Commission

## 1. VEP breakfast seminar

On 21 February 2014, the Flemish Evaluation Society (VEP) in Belgium together with the Department of Public Governance and the Research centre of the Flemish Government organized a breakfast discussion seminar on the draft evaluation policy guidelines of the European Commission.

The Flemish Evaluation Platform (Vlaams Evaluatieplatform - VEP) is an open network oriented at organizing and promoting the exchange of ideas, experiences, knowledge and information with regard to policy evaluation between the relevant actors (government, academic world, nonprofit organizations, advisory bodies, foundations, think tanks and consultancy) within all policy sectors and governmental levels. The VEP aims to build, strengthen and diffuse an evaluation culture, to enhance the evaluation capacity of the different actors, to rise the quality of policy evaluation and to stimulate the influence and use of policy evaluation. The VEP also aspires to be the contact point for international exchange within the field of evaluation theory and practice.

The session was chaired by Geert Steurs (Idea Consult and member of the board of VEP). Ms. Sara Piller (SecGen EU Commission) presented the background and contents of the Communication of the Commission and the draft guidelines<sup>1</sup>. Peter van Humbeeck (Social-Economic Council of Flanders, University Antwerp and member of the board of VEP) was asked to give a brief reflection on the draft evaluation policy guidelines of the European Commission and to raise some points for discussion.

This note summarizes the main points that came up in the presentations and discussion.

## 2. Introduction by Ms. Sara Piller

In its Communication on Smart Regulation (2010) the Commission wanted to target the whole policy cycle by attaching more importance to retrospective evaluation. This is a shared responsibility with the Member States, the European Parliament, the Council. Evaluation should also strengthen the voice of citizens and stakeholders. The communication introduced "Fitness Checks" and announced revision of the evaluation guidelines.

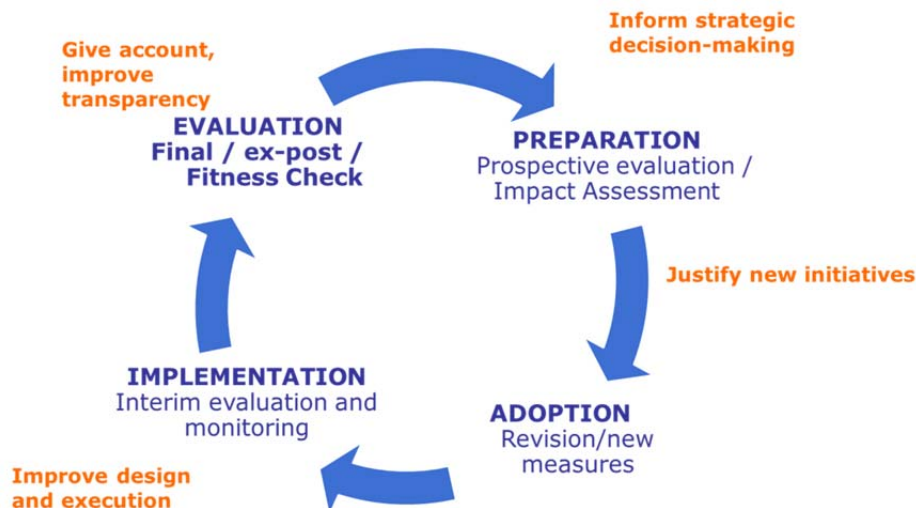
Ms. Piller explained further in detail the position of evaluation in the broader policy cycle (see figure) and the main purposes of evaluation. These are:

- Transparency, accountability. Society has a right to ask the administration to give account of what was done and achieved, so as to reduce asymmetries in information.
- Internal, organizational learning to improve the quality of an ongoing intervention and prepare the next one. It provides managerial information for those running the program or

<sup>1</sup> For more information, see [http://ec.europa.eu/smart-regulation/evaluation/consultation/index\\_en.htm](http://ec.europa.eu/smart-regulation/evaluation/consultation/index_en.htm) and <http://www.evaluatieplatform.be/programma/20140221.htm>

activity, i.e.: are we on the right track? Is the intervention still relevant , effective and efficient?

- Efficient resource allocation between interventions, between the separate elements of a specific program or activity, or between activities.
- Last but not least: inform decision-making, input to political priority-setting. Evaluation is a decision-making tool. It does not replace, but supports decision-making, both at a strategic (planning) level, and at the level of the design of a new intervention. It aims to raise the quality of the debate, taking into account the principles of Better Regulation and administrative simplification.



The Commission's attention for evaluation is not new. The key issue in Barroso's Commission 1 was Impact Assessment (IA); "Barroso 2" has looked at evaluation (from Better – to Smart Regulation). In his political guidelines (3 September 2009), President Barroso announced that: "We need to match this huge investment in ex ante assessment with an equivalent effort in ex post evaluation – to ensure that our proposals really do deliver what they promise and to enable us to revise and correct them where they fail to work as expected." When presenting his political program to the EP in September 2009, the President furthermore stated that he intended to "place ex post evaluation directly under his responsibility, to reflect the importance he gives to it". Similarly, the inclusion of explicit evaluation requirements in the EU Treaties, following the entry into force of the Lisbon Treaty, also reflects the growing importance which is being attached to retrospective evaluation.

Compared to earlier initiatives for evaluation, key element of Smart Regulation are a more timely input to decision making (learning from past *before* going forward), to evaluate all activities addressed to 3rd parties and a stronger focus on regulatory measures.

The Commission proposes a new definition of retrospective evaluation: « a critical, evidence-based judgment of whether an intervention has met the needs it aimed to satisfy and actually achieved its expected effects. It goes beyond an assessment of whether something happened or not, and looks at causality – whether the action taken by a given party altered behaviors and led to the expected changes and/or any other unintended or unexpected changes » (see Commission Communication COM (2013) 686).

There are four main reasons why the commission proposes new guidelines:

- To embed "evaluate first"; make sure evaluation answers IA and vice versa – complete the policy cycle/SR cycle. There can be a tendency to look forward and focus on new initiatives.

But changes are costly and take time to implement – so they need to be justified and greater attention needs to be paid to looking back before moving forward. There is a need to confirm the place of evaluation in the framework for EU action, to commit the appropriate resources and make sure that evaluations are conducted before proposing further action or change.

- To go beyond the state of play. Evaluations often present the state of play, they do not always include sufficient analysis of why something has happened, whether some of the change can be attributed to EU action and the extent to which the change fulfilled all initial expectations. Despite the commitment to evidence-based policy making, evaluations are not always supported by sufficient data and stakeholder opinion.
- To raise consistency in a decentralised system. Over time, there have been both consistent and diverging trends in Directorates General evaluation practices as they have evolved to accommodate new demands and developments, as well as the nature of different policy areas. This has led to differences in the planning, scope, method, content and final quality of evaluations. Not all EU legislation contains a clear commitment to conduct future retrospective evaluation, realistically scheduled so that the necessary data and information should be available.
- To build a clear and common scrutiny process to ensure high standard, critical approach and independent. At present, quality assurance is undertaken by the Directorate General's evaluation function and/or the steering group. Sometimes Directorates General draw on a wider group of resources, for example, by involving external stakeholders in steering groups or by involving academics in the quality assurance of a given evaluation. According to the evaluation standards, the steering group jointly produces (together with the officials managing the evaluation) a quality assessment of the final evaluation report. This quality assessment should critically assess the evaluation process undertaken, the information sources used, the analysis produced and the conclusions drawn. The quality assessment is not always published, in contrast to the IAB opinions, which are systematically made available alongside the final impact assessment and any associated Commission proposal.

The new guidelines are intended for internal use by the Commission services, providing an introduction to evaluation. The underlying principles can be summarized as:

- Encouraging those who are interested to get involved in evaluation. The guidelines suggest standard steps intended to make it easier for anyone interested: to know what is being evaluated and when the evaluation is taking place, to provide their input, to see all consultation results and a summary assessment of them and to understand whether action may follow-on from the evaluation. This would involve the central publication of:
  - 5yr evaluation plan providing a clear and comprehensive overview of all planned evaluations;
  - Evaluation mandates, published early in the process,;
  - The final evaluation report together with an assessment of its overall quality.
- Improving the coverage and usefulness of evaluation: seek to improve the strength and consistency of evaluations by:
  - Making sure that all evaluations address key questions on the performance and continued need for EU action. 5 minimum criteria,(effectiveness, efficiency, relevance, coherence and EU added value) or provide good reasons to justify why any of these elements has not been evaluated.
  - Defining the essential requirements for the conduct of an evaluation, distinguishing it from studies or reports containing only some evaluative elements.
  - Ensuring good advance planning and organisation of the collection of relevant data as well as good organisation of consultation of stakeholders and the general public.
- Improving co-ordination within the Commission, bringing together the different policy makers involved at an early stage and facilitating coherence. These actions should further strengthen

the role of evaluation and ensure that inputs are fully utilised in evaluations, improving the evidence base and providing more useful results. In turn, this would fulfil the commitment to "evaluate first" and complete the Smart Regulation cycle – evaluation results would systematically feed into any later impact assessment of policy options where change is envisaged.

The Commission welcomes all views and comments on how EU actions should be evaluated, based on the experience of EU evaluation, the experience in a national/local setting, and general thoughts and opinions. The public consultation is open until 25.2.2014.

### 3. Comments by Peter Van Humbeeck

Mr. Peter Van Humbeeck welcomed the new (draft) guidelines and the public consultation the commission is organizing.

#### What are the guidelines (not)

The Commission Guidelines for Evaluation are (draft) *evaluation policy* guidelines, and not *policy evaluation* guidelines. The focus is on the governance of evaluation, with measures to raise capacity and strengthen policy evaluation, such as requirements for planning, resources, transparency, due process, roles and division of tasks, quality assurance. It is not an users guide or resource book on how to perform evaluations.

In terms of the different levels of an evaluation system, the guidelines are situated at organizational level (arrangements and structures) and inter-organizational level (inter DG collaboration, working together with parliament, member states...) and nota at individual level (skills and competence) or societal level (values, culture, learning community of practice in broad civil society). Although there are important side steps towards the individual and societal level (annexes, training, stakeholder consultation...).

This seems to be the right choice since there are already numerous resource books available. What is needed, are (new) governance agreements and arrangements.

The guidelines furthermore are inspiring for member states and evaluation societies like VEP. It raises several important issues and contains many useful ideas. In Flanders, some actors (SERV, SBOV) already have made recommendations to strengthen policy evaluation after the coming elections. The ideas and proposals of the Commission are very useful to deepen those recommendations and perhaps take an initiative with VEP towards the new government.

#### Some strong points

Mr. Peter Van Humbeeck summarized some strong points of the guidelines in 10 bullets.

- The communication and guidelines propose a substantial, ambitious program for evaluation, that is being institutionalized.
- The proposed organizational roles and structures seem to be well founded: a decentralized system with central support and guidance, minimum standards, role of the steering group...
- The guidelines show a careful balance between general guidelines and space for tailoring an evaluation to the specific needs and circumstances, because how to evaluate depends on what to evaluate (cf. minimum requirements with further recommendations...)

The guidelines have many good or even excellent points and ideas. Specific points Peter Van Humbeeck stressed are:

- The (new) definition and purpose of evaluations (and the essential requirement to distinguish evaluations from studies with evaluation components).
- The broad scope (all measures and actions, not only financial programs, even not only legislation).
- The strong focus on utilization of results: the attention for end use, provisions to stimulate use, but also discussion and debate; to improve the use of evaluation in policy preparation (closing the policy cycle, evaluate first principle, linked with IA, identify follow up actions within six months); This reflects a real need: often new measures are introduced without looking into what is wrong with the existing or why they do not work as expected, adding to complexity and ineffective and inefficient regulation.
- The (new) instruments for high quality evaluations: to promote evidence-based, robust conclusions, advance planning and organization; independence, mandatory quality assessment...
- The transparency of final results and conclusions, but also of what is being evaluated, when it is taking place, how is done and what is being done afterwards: see evaluation plan (coupled with existing planning tools), evaluation mandates for each individual evaluation, transparency, accountability and stakeholder participation, informing the debate...
- The five mandatory evaluation criteria (with the necessary flexibility). A particular strong point is that the criteria focus on needs (relevance) beyond effectiveness (the problems behind the objectives, end user perspective; outcomes, impacts beyond effects...). Also positive is that “effectiveness” is more than assessing what has happened: what caused change, how much of this can be attributed to EU action? Very positive is that “efficiency” is mandatory. This is a difficult, important one that is often neglected in evaluations. But evaluating efficiency is not easy: it supposes a comparison of costs with benefits, or of cost of the policy with costs of alternative policies or policy instruments. The (dangerous) risk is that “efficiency” will therefore be narrowed to administrative burdens, SME-costs ... Coherence is also very important. Cf. e.g.
- renewable energy support scheme in Flanders: it is successful in reaching short term RE targets (effectiveness), but at a very high cost (over subsidization, windfall profits leading to a low efficiency) and with very negative effects for long term market integration and competition in electricity markets (coherence).
- The attention for learning, not only on the policy that is being evaluated, but also at meta level (cross DG quality review panel (meta evaluation), role SG...).

### Comments and questions

Peter Van Humbeeck also raised 10 comments and questions.

- Open process? The importance of buy-in of stakeholders and stakeholder input is recognized, but the impression sometimes is that in the end the focus is on top down transparency to enhance quality and accountability rather than on real bottom up participation, active involvement and real engagement. Transparency in all steps in the process is very positive, but can or should it not be more? To get involved, more is necessary than transparency of the publication of the forward planning, evaluation mandates, final reports and assessments of their quality. Peter Van Humbeeck illustrated his point with three examples:
  - What about stakeholder participation in the screening of what to evaluate first? This seems to be absent. The focus is on governments as a source of demand for evaluations, not on civil society as a source for demand for evaluation.
  - The draft guidelines unclear about content of the evaluation mandate. De Communication speaks about purpose, scope, steps in the process, when consultation takes place... It says that stakeholders should know from an early point in the process when they will have the opportunity to participate. The draft guidelines seem to limit the



evaluation mandate to: “purpose and objectives, key questions, scope, expected outputs, deadlines, quality criteria”. Stakeholder consultation seems to be forgotten.

- The draft guidelines speak about publication of evaluation reports, but not about circulation of draft reports and findings for comments or feedback on preliminary findings. This is however necessary to engage stakeholders, not to influence but to challenge the conclusions and recommendations.
- Quality of stakeholder consultations? The guidelines refer to the EU minimum requirement. The practice is often line consultation. This is not enough. Good consultation requires further guidance, careful planning and adequate selection of consultation methods.
- The ambition and screening phase? This looks a black spot in de guidelines: who, when, how are decisions made on what to evaluate first? Is there involvement of stakeholders? How are priorities being set? Furthermore, there seems to be an imbalance between the ambition in the communication to evaluate first (“all significant proposals for new or revised legislation are in principle based on an evaluation of what is already in place”) and the ambition in the draft guidelines to “conduct one evaluation each year in each DG”.
- The evaluation criteria? The mandatory criteria are important (cf. supra) but distribution of cost and benefits, equity and more general social issues seem to be neglected somewhat.
- The components of a final evaluation report? According to the draft guidelines it should contain conclusions, but what about recommendations? Should an evaluation not conclude with recommendations?
- Expertise? A desk officer will not be expected to become an expert in evaluation methodology, because such expertise is normally easily available from external consultants. But in house expertise will still be necessary through training, guidance (annexes?)
- The policy towards member states? Success also depends on the capacity and culture in member states? What are the plans or efforts of the commission? In some member states there is almost no beginning of an evaluation policy... One example is the implementation plan for member states for EU Directives: they should e.g. go into requirements for data collection and evaluation because thinking about and organizing data collection needs to be established from the start of the action/part of the implementation. How will joint evaluations with member states be organized?
- Regulatory fitness initiative? This needs further clarification, notably regarding the link with the evaluation policy.
- The of role the SG? What is the role of the SG in monitoring compliance? When, How?
- The public consultation of the draft guidelines? Have there been many comments? What are the main issues?

#### 4. Discussion

Following the points raised by ms. Piller and mr. Van Humbeeck, the participants formulated some additional issues.

- How does the commission cope with the differences between member states regarding data, evaluation methods, capability...?
- How large is the commitment of the political level (council, commissioners)? (cf. experience in Belgium and Flanders: there is commitment for IA on paper, but actual commitment is low resulting in few (high quality) IA and evaluations.
- What is the role of the European Parliament in the evaluation policy? The guidelines do not contain much information on this.
- What is the role of the Member states? How to improve their participation in an early phase?
- What is the role of Sec Gen? What is the experience? Does the decentralized system with oversight from Sec Gen work? (and can this be an example for Flanders?)

- What if the political decision of the commission deviates from the preferred option in a IA, as sometimes is the case (e.g. shale gas, invasive species, ...)? Should it not be more motivated in a transparent way? Without a clear justification, evaluation becomes difficult.
- Should there not be more attention for social criteria and social effects in the commission guidelines?
- The new definition of policy evaluation puts more emphasis on causality. Will there be more guidance for attribution/contribution analysis?
- How is the important issue of independence in policy evaluations organized?
- How were the programs selected for Refit-Top Ten? What were the criteria? How were priorities set?
- Why is “better regulation” (only) on the agenda of the competitiveness council and not on the agenda of other council meetings?

Almost all questions and comments concerned the proposed Commission evaluation approach. Flemish evaluation experiences and exercises were hardly raised.

Ms. Piller thanked all participants for the comments and went through most of the issues that were raised, explaining the context and rationale behind the proposal of the commission.

She would welcome a written input to integrate it into the public consultation and the feedback of the commission.

VEP promised to do so by means of this note.

## 5. Concluding suggestions

VEP welcomes the draft Commission Guidelines for Evaluation. It considers this initiative as an important step towards improved policy making at the EU-level.

VEP would like to stress that the future Commission evaluation policy - in order to realize its full potential as elaborated in its four main purposes - needs to be applied in a collaborative way. The draft Commission Guidelines address this issue but remain rather vague.

Given the role of the authorities (national, regional, local) in the member states to implement EU policies, it is obvious that the future Commission evaluation policy should address this issue more precisely.

As a means to initiate the communication between the Commission and the member states, we suggest the Secretariat-General to establish an EU policy evaluation network or platform with experts/representatives from the member states and the Commission. This network could focus on information exchange, mutual learning about good practices, awareness raising and capacity improvement.

It might also operate as an additional information channel within the context/scope of particular Commission evaluation exercises.

The long term ambition of this network should be its contribution to co-create a joint EU policy evaluation culture as an element of a collaborative government model.<sup>2</sup>

<sup>2</sup> See, Commission, DG Connect, A vision for public services, draft version dated 13/06/2013, <http://ec.europa.eu/digital-agenda/en/news/vision-public-services>