

Evaluation of European Commission Policy and Programmes. Overarching Principles and Quality Assurance

Joan Ribí

Smart Regulation, developed during the last three years reflects a shift in emphasis to people's well-being and progress. It gives emphasis to stakeholder consultations, to strengthen the voice of citizens and stakeholders in the policy making process. This puts into practice the provisions of the Lisbon Treaty on participatory democracy. Smart Regulation increases transparency and accountability and promotes evidence-based policy making. In this paper I refer to the standards of four major programmes and reflect on how quality is embedded in policy, programmes and projects of the European Commission through Evaluation, synthesising recent developments.

Table of Contents

- 1. Introduction**
- 2. The Revised Framework**
- 3. Smart Regulation in the European Union: managing the quality of regulation throughout the policy cycle**
- 4. Overarching Principles**
- 5. Quality assurance and Quality control**
- 6. Four major programmes**
- 7. Ex ante evaluations of programmes**
- 8. Ex-post Evaluations**

1. Introduction

All projects and programmes funded by the European Commission undergo ex-ante and ex-post evaluation. The evaluation cycle comprises: Ex-ante (Feasibility) -> Ongoing/mid term -> Ex-post/Results Evaluations. Ex-ante documents needs to be addressed, the results to be obtained and the feasibility of planned projects; interim evaluations document progress and implementation of the intervention; ex-post focuses on results and the contribution of the programme to change. All ex-ante and retrospective evaluations of all EU interventions must comply with Commission Evaluation Standards.

Ex-ante evaluation examines the potential risks for the programme, both in relation to policy choices and the implementation system proposed (EC Impact assessment Guidelines 2009). Ex-ante evaluations are used for a range of activities such as potential impact assessment of legislation or policy, can address a piece of legislation, a policy, a programme or a project. Ex-ante evaluations improve quality, relevance and comprehensiveness of design of new programmes and provide information on the basis of which decision makers can judge the

value of the programme proposal. Ex-ante evaluation of projects still to be implemented focus on the quality of project deliverables and likely impact of the intervention. Evaluation principles embed quality to ensure only high quality projects are funded and appropriate project results are delivered.

Retrospective evaluation (interim, ex-post and final evaluation) uses impact assessment to provide insights for improving policy and legislation.

2. The Revised Framework

Traditionally evaluations have examined individual funding programmes or pieces of legislation. Attention is now paid to evaluation of broad areas of legislation and cross-cutting issues. With a view to improving the political relevance a revised framework for policy and programme evaluation is set out. All retrospective evaluations (Fitness Checks for policy, final, ex post and interim) of all EU interventions (actions) are guided by Smart Regulation (COM 2010, 543 final, COM 2012, 746 final, COM 2013, 686 final) which also stipulate ex-ante (Impact assessment). A guideline for Smart Regulation applicable to all retrospective evaluations is being developed subsequent to a Public consultation (Public consultation on Commission Guidelines for (retrospective) evaluation November 2013). A financial regulation requires ex-ante and ex-post evaluation of all EU funded programmes.

3. Smart Regulation in the European Union: managing the quality of regulation throughout the policy cycle.

Smart Regulation addresses the whole policy cycle- from the design of a piece of legislation to implementation, monitoring, evaluation and revision. New legislation is built on the strengths of an impact assessment system. Smart regulation is based on the shared responsibility of European institutions and of Member States, because it is an integral part of collective efforts in all policy areas. In this respect, consultations are used to strengthen the voice of citizens and stakeholders. This puts into practice the provisions of the Lisbon Treaty on participatory democracy. Smart Regulation aims to design and deliver regulation that respects the principles of subsidiarity and proportionality attaching great importance to evaluating the functioning and effectiveness of existing legislation.

Ex-post evaluation of legislation evaluates the effectiveness and efficiency of EU legislation to improve the quality of policy making, identify opportunities to simplify legislation and reduce administrative burdens. The principle `evaluate first`, a key component of Smart Regulation, ensures that all significant proposals for new or revised legislation are based on an evaluation of what is already in place.

Individual Policies

`Fitness checks` of individual policies assess if the regulatory framework of a policy area is fit for purpose and if not, what should be changed. This is extended by the REFIT Programme (Regulatory Fitness and Performance Programme) implemented in a transparent manner. A system of impact assessment prepares evidence for political decision-making and to improve transparency on the benefits and costs of policy choices. Hence measures proposed by the Commission will be necessary, cost effective and of high quality.

Smart Regulation Policy deploys the tools: (i) Impact assessment, transparently assessing legislative and non-legislative policy options by comparing both potential benefits and costs

in economic, social and environmental terms (ii) Evaluation of EU Regulatory measures, including evaluation of financial programmes.

Evaluation is a key tool within Smart Regulation providing a critical, evidence – based judgement of whether an intervention has met the needs it aimed to satisfy and actually achieved its expected effects and looks for causality – whether the action taken by a given party altered behaviours and led to the expected changes. Evaluation assesses what has happened, why it occurred and how much has changed as a consequence. Evaluations must be independent, impartial, evidence based so that causality is based on facts and real-life chain of events.

Smart Regulation provides a high quality assurance within the evaluation system through: transparency, clear rules for conducting evaluations (evaluation standards), a Steering Committee, Quality assessment of the final report, independent evaluators, a quality review panel to comment on the quality of finalized evaluations and identify strengths, weaknesses and lessons learnt.

4.Overarching Principles

The Overarching Principles in all ex-ante and retrospective evaluation at any stage in the Project Cycle are embedded in the judgement criteria: relevance, effectiveness, efficiency, coherence and EU added value. All evaluations (ex-ante and ex-post) must address these mandatory evaluation criteria, both in the case of external sourcing and when Commission staff undertake evaluations internally.

- Relevance to policy and the Call objectives;
- Effectiveness of intervention in causing changes/meeting the objectives;
- Efficiency involves justification of costs and achievements, (ratio of costs to benefits);
- Coherence with other interventions with similar objective;
- Added value of EU intervention compared to what can be achieved by Member States at national and regional level;
- Others are added depending on the programme, such as sustainability (the duration of changes caused by the project after end of project), utility (the extent to which changes/effects of an intervention satisfy or not stakeholder`s need; the degree of satisfaction of the various stakeholders), complementarity (the extent to which the intervention supports other policies, both EU and Member States), Coordination (the extent to which interventions are organised to maximize their joint effects ie mobilizing resources combined with harmonizing measures), Equity (how fairly the different effects are distributed across the different stakeholders, regions, genders), acceptability (positive and negative perceptions to the intervention of targeted stakeholders and the public).

5.Quality assurance and Quality control

Quality can be considered as a characteristic of the process (planning and design) and as a characteristic of the product (evaluation) embedded in Evaluation Standards.

Evaluation standards set out the governing principles which provide a basis for quality assurance along all the evaluation process are required by Smart Regulation. Smart

Regulation Evaluation Standards require that Evaluation activities are (A) appropriately organised and resourced to meet their purposes (B) planned in a transparent and consistent way so that the relevant evaluation results are available in due time for operational and strategic decision-making and reporting needs (C) Evaluation design must provide clear and specific objectives, and appropriate methods and means for managing the evaluation process and its results (D) Evaluation activities must be conducted to provide reliable, robust and complete results (E) Evaluation results must be published and followed up to ensure maximum use of the results and that they meet the needs of decision makers and stakeholders (F) have a minimum list of quality criteria.

Process Criteria for Quality Assurance:

- Coherent and specific objectives
- Well drawn terms of reference
- Sound tender selection process
- Effective dialogue & feedback thro evaluation
- adequate information resources available
- Good management and coordination by evaluation team
- Effective dissemination of reports to Steering Committee and policy/programme managers
- Effective dissemination to decision-makers and stakeholders

Output Criteria for Quality Control- The Quality assessment of the final evaluation

report is assessed on the basis of pre-established criteria. The minimum criteria for assessing the quality of an evaluation report are:

- Meeting needs as laid out in ToR
- Relevant scope and coverage
- Open process-partners and stakeholders involved in the evaluation design and discussion of results
- Defensible design and methods
- Assumptions and limitations clearly explained
- Reliable data used
- Sound data analysis. Data sufficient and triangulated
- Credible results that relate to analysis and data
- End user clearly addressed
- Impartial conclusions showing no bias and demonstrating sound judgement
- Useful recommendations
- There is a clear and logical chain between evaluation questions, evidence, findings, conclusions, recommendations
- Clear report with executive summaries and annexed supportive data

6. Four major programmes

In this paper I review evaluation in the following Programmes to discuss how overarching principles are embedded in evaluation.

- **HORIZON 2020**, the EU Research and Innovation programme, the financial instrument implementing the Innovation Union, a Europe 2020 flagship initiative

aimed at securing Europe`s global competitiveness, is established by Regulation No 1291/2013 which lays down the specific rules for participation.

- **EU Regional Policy for Socio-Economic Development: Cohesion Policy**, established by the Common Provisions Regulation (CPR) Regulation (EU) No 1303/2013 and others is an investment policy supporting the delivery of Europe 2020 Strategy: job creation, competitiveness, economic growth, improved quality of life, sustainable development and implemented through Structural Funds (made up of the European Regional Development Fund (ERDF) and the European Social Fund (ESF)). The Structural Funds and the Cohesion Fund make up one of the largest items of the budget of the European Union.
- **Evaluation of State Aid** implemented through international agreements with country governments.
- **EuropeAid** Development and Cooperation, aims at reducing poverty in the world, ensuring sustainable development, promoting democracy, peace and security; The main areas are: Governance, Democracy, Human Rights and Support for Economic and Structural Reforms (EuropeAid Mission, Version 18/11/2013).

7. Ex ante evaluations of programmes

These identify:

- The need to be met in the short or long term;
- The objectives to be realised;
- The results expected and the indicators needed to measure them;
- The added value of Community involvement;
- The risks linked with proposals and alternative options available;
- The lessons learnt from similar experiences in the past;
- The human resources and other administrative expenditure to be allocated with due regard to the cost effectiveness principle;
- The monitoring system to be set up.

Example: Ex-ante evaluation of H2020 projects. The quality of H2020 projects is evaluated under three specific award criteria: Excellence, Impact, Quality and Efficiency of Implementation.

- **Excellence.** Clarity and pertinence of objectives; Credibility of proposed approach. Soundness of concept; Progress beyond current situation;
- **Impact.** The extent to which the outputs of the project contributes to the expected impact; Negative/positive environment impacts; Ethical issues; Time frame and budget.
- **Quality and efficiency of the implementation.** The assumptions and external factors that may determine whether these impacts will be achieved; The end users (who are major stakeholders able to exploit the results) are part of the consortium thus increasing impact; Strategic Plan; Coherence and effectiveness of the work plan, including appropriateness of the of the allocation of tasks and resources; Appropriateness of the management structures and procedures including risk assessment; Monitoring and indicators of progress of project planned.

8. Ex-post Evaluations

Ex-post evaluation questions, defined in the start-up of any evaluation can be at different levels. They can be:

- Descriptive questions intended to observe, describe and measure changes (what happened?)
- Causal questions which strive to understand and assess relations of cause and effect (how and to what extent is that which occurred attributable to the intervention?)
- Normative questions which apply evaluation criteria (are the results satisfactory in relation to targets?)
- Predictive questions, which attempt to anticipate what will happen as a result of planned interventions (will the interventions to counter unemployment in this territory create negative effects for the environment or existing employers?)
- Critical questions, which are intended to support change often from value-committed stance (how can equal opportunity policies be better accepted by SMEs? or what are the effective strategies to reduce social exclusion?)
- Ideally, evaluation questions should have the following qualities:
- The question must correspond to a real need for information, understanding or identification of new solution. If a question is only of interest in terms of new knowledge, without an immediate input into decision-making or public debate, it is more a matter of scientific research and should not be included in an evaluation.
- The question concerns a result or a need. That is to say, it concerns, at least partly, elements outside the programme, notably its beneficiaries or its economic and social environment. If a question concerns only the internal management of resources and outputs, it can probably be treated more efficiently in the course of monitoring or audit.
- The question concerns only one judgement criterion. This quality of an evaluation question may sometimes be difficult to achieve, but experience has shown that it is a key factor in the usefulness of the evaluation. Without judgement criteria clearly stated from the outset, the evaluation rarely provides conclusions.

Example: Evaluation of EU Regional Policy for Socio-Economic Development (DG REGIO)

The Structural and Cohesion funds are organised in programmes and evaluation takes place at the ex-ante, interim and ex-post stages. Programmes often form just one input to achieving wider policy goals and their evaluation contributes to policy evaluation. The programmes comprise many interventions and projects. Evaluation can take place at the level of the priority action, project or programme level. Socio-economic development undertakes policy, project and thematic levels evaluations. Ex-ante evaluation (which must contain an environmental impact assessment) is the responsibility of Member States and regions preparing the programmes. The ex-ante evaluation is forwarded to the European Commission with the proposed operational programme, but the European Commission has no role in assessing or approving this evaluation. A financial regulation requires ex-ante and ex-post evaluation of all EU funded programmes. The mid-term evaluation is undertaken by an independent evaluator.

Ex-ante criteria. Ex ante evaluation examines the potential risks for the programme, both in relation to policy choices and the implementation system proposed. The programme developers develop detailed questions to be answered in relation to the national, regional or sectorial strategies to be evaluated. Evaluation criteria are formulated as questions relating to

the judgement criteria: relevance, effectiveness, efficiency, sustainability and utility. The Strategy is based on Specific objectives and intended result, targeted and achieved outputs and impact (actual result). An ex-ante evaluation shall incorporate where appropriate the requirements of an SEA (article 48(4)) Common Provision Regulation.

Ex-post evaluation. Programme theory, the elaboration of Logic Models, used extensively in the context of World Bank and EU funded development programmes is a simple evaluation theory that focuses on programme inputs, outputs, results and impacts. The theory of Change, now being pursued, goes beyond input output descriptions, seeking to understand the theories of actors with regard to programme interventions and why they should work. The new logic shifts focus from actions and financial means to their outcome in terms of people's well-being and progress through appropriate indicators within an effective monitoring and evaluation system.

Two broad categories of impact evaluation are recognised:

- Theory based impact evaluation follows each step of the intervention logic identifying links and mechanisms of change (how and why) accompanied by hypothesis, answering the questions why an intervention works or don't and under what circumstances. This approach mainly produces a qualitative estimate of the impacts, a narrative. It is based on approaches such as Process tracing, Contribution analysis, General elimination method and includes methods such as literature review, text analysis, interviews, surveys, focus groups and case studies.
- Counterfactual impact evaluation (causal inference) is used to understand whether the intervention works and how much difference it makes. Evaluations of this type are based on models of cause and effect. This method is useful in answering how much of the change is due to the intervention and for whom. It thus requires a control or comparison group, a large enough number of participants and entities and controls for statistical significance and good data on supported and non-supported participants and entities to compare results. A good evaluation plan is required from the outset.
- Ideally counterfactual and theory based approaches complement each other.

Inputs, targeted outputs, achieved outputs, outcomes

The concept that is central to the new framework: intended outcome is defined as the specific dimension of the well-being and progress of people (in their capacity of consumers, workers, entrepreneurs, family and community members, etc.) that motivates policy action ie that is expected to be modified by the interventions designed and implemented by a policy. Example of outcomes:

- competence (the increase of which is the aim of providing enhanced education/child care/implementing urban renewal);

Outputs are defined as measurable policy actions whose intended task is to produce actions. In this example, enhanced education/child care/implementing urban renewal is an output: measurable policy actions whose intended task is to produce outcomes.

The chosen outcome must be represented by appropriate measures by selecting appropriate outcome indicators, their baselines and their targets,

Once outcome indicators are chosen, a baseline must be established, ie the value of the outcome indicator before the new policy intervention at stake is undertaken and the expected trend. Outcome targets are established, ie the values of the indicators which policy action aims to achieve. The target reflects the theory of change with regards to policy.

Evaluation of State Aid

The overall objective of State aid ex-post evaluation is the verification of the balance between the public objective of the aid and its impact on competition and trade between Member States. State aid evaluation can explain whether and to what extent the original objectives of an aid measure have been fulfilled (i.e. assessing the positive effects) and to determine the impact on markets and competition (i.e. possible negative effects). State aid evaluation in particular allows:

- Assessing whether the scheme is effective in achieving the direct objective for which it was introduced (e.g. whether a certain aid has effectively led to more investment on the part of the aid beneficiaries);
- Assessing whether the scheme provides for indirect effects on the objective of interest (e.g. whether the aid led to positive spill-over effects on other firms, or led competitors to lower their own investment effort);
- Detecting negative indirect effects, in particular the potential aggregated effect of large schemes on competition and trade (e.g. whether the aid led to crowding out or shrinkage of competitors or just displaced activity away from other locations, against the EU interest).

Based on the above assessment, evaluation can confirm whether the assumptions underlying the ex-ante approval of the aid scheme are still valid and/or help to improve the design of the scheme, introduce corrective measures, and calibrate State interventions to improve the effectiveness and efficiency of the aid to the extent necessary to guarantee that the overall balance between positive effects and distortions is still acceptable

Evaluation of EuropeAid Programmes

Evaluation aims at rendering accounts to the public on the results and impacts of activities financed by EC funds and drawing lessons on what has worked and what has not. In addition to the listed Evaluation Criteria (judgement criteria), there is an additional criterion, impact: positive, negative, primary and secondary, directly and indirectly, intended or unintended;

Nine Quality Criteria are used by the evaluation manager to assess the quality of the final report. They are not used in the context of evaluation or judgement criteria.

- Meeting needs: does the report adequately address the information needs of the commissioning body and fit the terms of reference?
- Relevant scope: is the rationale of the intervention and its set of outputs, results and impacts examined fully, including both intended and unexpected policy interactions and consequences?
- Defensible design: is the evaluation design appropriate and adequate to ensure that the full set of findings, along with methodological limitations, is made accessible for answering the main evaluation questions?
- Reliable data: are the primary and secondary data selected adequate? Are they sufficiently reliable for their intended use? When statistical data are missing, are qualitative data meaningful? When only weak data are available, the evaluators have explained their weaknesses and the limit of use.
- Sound analysis: is the analysis of quantitative and qualitative data appropriately and systematically done so that evaluation questions and judgement criteria are informed in a valid way? Are cause and effect links between the intervention and its results explained? Are external factors correctly taken into consideration? Are comparisons made explicit?

- Credible findings: do findings follow logically from and are justified by the data analysis? Are findings based on carefully described assumptions and rationale?
- Valid conclusions: are conclusions linked to the findings? Are conclusions based on the judgement criteria? Are conclusions clear, clustered and prioritised?
- Useful recommendations: are the recommendations linked to the conclusions? Are they fair, unbiased by personal or stakeholders' views, and sufficiently detailed to be operationally applicable? Are they clustered and prioritised?
- Clear report: does the report describe the intervention being evaluated, including its context and purpose together with the process and findings of the evaluation? Is the report easy to read and has a short but comprehensive summary? Does the report contain graphs and tables.

EuropeAid supports the BUSAN international principles of effective aid and good development. These principles include: country leadership and ownership of development strategies; a focus on results that matter to the poor in developing countries; inclusive partnerships among development actors based on mutual trust; and transparency and accountability to one another.

Dr. Joan Ribí, B.Sc Hons.UK, M.Sc UK, PhD Australia, Comp.Inst Eng.Australia

References

- Smart Regulation (COM 2010, 543 final, Smart Regulation in the European Union.
- COM 2012, 746 final, EU Regulatory fitness.
- COM 2013, 686 final, Strengthening the foundations of SMART Regulation-improving evaluation.
- Public consultation on Commission Guidelines for evaluation November 2013.
- EVALSED: The resource for the evaluation of Socio-Economic Development, November 2013.
- European Commission, Impact Assessment Guidelines 15 January 2009, consultation to be launched, July 2014.
- DG for Development and Cooperation – EuropeAid Mission, Version 18/11/2013.
- Evaluation Methods for the European Union External assistance. Guidelines for project and programme evaluation Volume 3.
- European Commission Evaluation in the field of State Aid – concepts and recommendations November 2013.