Expert Evaluation Network

Expert review of selected evaluations and studies

Note on the provisional results of the peer group exercise (Task 4)

Core team

Context

In the current programming period 2007-2013, Member States have a high degree of discretion and flexibility on how they exercise their responsibility to evaluate interventions supported under Cohesion policy.

In the present programming period, there is significant variation between Member States in the number and nature of evaluations carried out on Cohesion policy programmes. This is partly a result of differences in the policies pursued and the pace of advancement of the implementation of programmes but equally of differences between Member States in evaluation culture and the capacity to plan, coordinate and carry out evaluations. The majority of the evaluations undertaken have been focussed on implementation and procedural aspects and on how to speed up financial absorption. Relatively few have been concerned with assessing the effect of interventions on policy objectives. However, it is encouraging that some of the advances in evaluation methods have begun to appear in Cohesion policy evaluations.

In order to encourage a move in this direction, DG Regional and Urban Policy established an expert review process at the beginning of 2011 as part of its Expert Evaluation Network. This process involved a group of leading evaluation experts being charged with examining selected evaluations. In all, 15 evaluations have so far been reviewed by the peer group (7 quantitative counterfactual and 8 based on qualitative or mixed approaches). This note summarises the main points to come out of the reviews and the discussions at the four meetings of the group which have taken place up to now.
(between June 2011 and December 2012). It indicates strengths and main deficiencies of the evaluations reviewed and the improvements that are necessary if a results-oriented policy is to be established. It suggests also the action that could be taken by DG Regional and Urban Policy in support of this.

**Main points from the peer reviews**

Although all the evaluations reviewed were selected because they seemed on initial inspection to be of comparatively high quality (i.e. in relation to others produced on ERDF-supported interventions), only a small number were assessed by the group as being of good standard. In most cases, deficiencies were identified either in the approach adopted or in the analysis undertaken or in both. The most common deficiencies, which match those identified in the literature, are:

- **an inadequate link between the evaluation questions and the design of the evaluation**: this is often due to the evaluation question being specified in too vague a way or not being clearly defined,

- **shortfalls in the evaluation design**: these derive, in particular, from a failure to adopt an appropriate methodology given the nature of the intervention concerned, insufficient attention to the context in which the intervention has taken place, a lack of technical competence on the part of the evaluators;

- **ineffective application of the methodology and techniques used**: questionnaires and surveys are sometimes not well enough designed to avoid biased responses, a triangulation with different techniques is rarely carried out and clear explanations of the techniques available and justifications for the choice of method(s) made are equally rare. Methodological shortcomings seem to be due not only to a lack of knowledge and understanding on the part of evaluators, but also to insufficient guidance and supervision from commissioning bodies.

- **the use of insufficient or inappropriate data**, which is essential if meaningful results are to be produced;

- **a lack of plausible connection between the evidence assembled and the conclusions drawn on the effects of the intervention**, due to weaknesses in the underlying assumptions made in the study or in its internal logic, a failure to understand the
rationale, or theory, underlying the intervention or a desire on the part of the evaluator to come up with findings which conform to the expectations, or wishes, of the client;

- a lack of conciseness of the evaluation reports produced and the inclusion of irrelevant or redundant content: in many cases the reports are excessively long and not sufficiently focused, which reduces their readability, but at the same time they often lack a clear explanation of what the intervention is intended to achieve, how it is intended to do so and how it operates in practice;

- poor informative value of the evaluation in respect of policy implications: in a number of cases, the evaluations reviewed contained many uninformative statements and failed to set out clearly the implications of the findings for current and future policy.

A major point is that the poor quality of the evaluations reviewed seems to be more recurrent when carried out on ERDF-supported interventions. It is not evident to the same extent in evaluations of national policies in particular areas, especially when carried out by academic researchers. In a number of countries, there is no shortage of evaluations of a reasonable standard and there are effective means of control over the methods used in operation (for example, as regards assessing the effect of employment policies in Germany).

The weaknesses in the evaluations reviewed, therefore, could in part to stem from insufficient time and resources dedicated to evaluation in comparison to efforts related to the implementation of programmes ("spending the money") resulting in ineffective evaluation procedures, such as the limited involvement of the right stakeholders in defining the evaluation questions, a lack of effective control by commissioning authorities, inadequate terms of reference; and the absence of a meaningful discussion of evaluation results and the limited extent to which these feed into policy. Note again that the review concerned a demanding type of evaluations - impact evaluations. An assessment of evaluations looking into implementation issues and readjusting programmes to a changing context would most likely a far more positive picture.
The way forward

A number of implications can be drawn from the experience so far as regards ways in which the situation can be improved.

- *Clear objectives in programmes* would facilitate better evaluations.
- *Improving terms of reference*: the development of guidelines for drafting ‘good’ terms of reference (in part through borrowing from the UN and other international organisations) including concrete examples with a list of “do’s and don’ts”; the possible organisation of a competition for best terms of references, encouraging the development of training for national officials (including through Summer schools); and the exchange of information and experience among commissioning authorities;
- *Increasing evaluation capacity in public authorities*: the development of guidelines for formulating an appropriate evaluation plan, the organisation of workshops and coaching of the officials responsible for evaluations in Managing Authorities and implementing agencies to improve their ability not only to commission evaluations but to assess both the proposals received and the evaluation reports produced.
- *Increasing the diffusion of sound evaluation methods*: the publication of good examples of evaluations (currently being done through the DG Regional Policy website); the publication (on the website) of the peer reviews produced by the exercise described here after notifying the authors of the studies concerned; the launch of small-scale evaluations with open terms of reference in order to assess the different approaches suggested to evaluate particular issues or policy measures.
- *Increasing access to data*: facilitating access to the data required to carry out evaluations through small grants and through funding prizes on evaluation studies, in cooperation with Member States which make the necessary data available. This would especially facilitate evaluations carried out by academic researchers.