

# Mid-Term Evaluation of Rural Development Plans in Italy: Comparing models<sup>1</sup>

VINCENZO FUCILLI\*

Jel classification: R110, R580

## 1. Introduction

The planning and implementation of community operations for rural areas, as outlined in regulations for Structural Funds (Reg. EC No. 1260/99) and for Rural Development (Reg. EC No. 1257/99), have been based on the Rural Development Plan (RDP). The Rural Development Plan, which has been an innovation with respect to the previous planning cycle, has financed all rural development operations covered by Community regulations in areas outside Objective 1 (Ob. 1). In the Ob. 1 regions, the RDP has contained only the accompanying measures (agro-environmental incentives, early retirement and forestation) and compensation for areas that are disadvantaged and subject to environmental constraints. The other measures included in Reg. EC 1257/99 are integrated in the 7 Italian Regional Operational Programmes. Therefore, the RDP has been the only single programming document which has comprehensively and systematically put rural development policy into practice. The autonomous Italian regions and provinces outside Ob. 1 have in fact opted for single planning of rural development operations. Furthermore, with the reform of

## Abstract

The paper comes from the observation that, although there are many evaluations of the rural development planning, there does not appear to be a similar number of studies and analyses of the «evaluation» itself and, in particular, studies and analyses of complex evaluations like those concerning the Rural Development Plans.

On the basis of the contents of RDPs mid-term evaluations, the level of coherence of the evaluative practice in Italy will be examined, in relation with the methodological debate that has grown on the evaluation of programs as opposed to the EU Commission methodological prescriptions. On the basis of (meta) criteria, the aim of the work is to verify the methodological contents of the mid-term evaluation reports of 14 RDPs in the areas of the central-northern Italy, by analysing the evaluation processes rather than their outcomes. Through analyses and comparisons of these reports, the «quality» of evaluations is reconstructed.

**Keywords:** Meta-evaluation, Evaluation, Rural Development Plans.

## Résumé

*L'idée d'écrire cet article dérive de l'observation du fait que même s'il y a de nombreuses évaluations sur la planification du développement rural, il n'y a pas assez d'études et analyses sur le processus d'évaluation lui-même et, en particulier, d'études et analyses sur les évaluations complexes comme celles qui concernent les Plans de Développement Rural (PDR). Sur la base des contenus des évaluations à moyen terme des Plans de Développement Rural, le niveau de cohérence de la pratique évaluative mis en œuvre en Italie sera examiné, en relation avec le débat méthodologique sur l'évaluation des programmes qui s'oppose aux prescriptions méthodologiques de la Commission européenne. Sur la base de (méta-)critères, cet article vise à vérifier les contenus méthodologiques des rapports d'évaluation à moyen terme de 14 plans de développement rural mis sur pied en Italie du Centre et du Nord, en analysant les processus d'évaluation plutôt que leur résultats. A travers l'analyse et la comparaison entre ces différents rapports, nous avons établi la « qualité » des évaluations faites.*

**Mots clés:** Méta-évaluation, Evaluation, Plans de Développement Rural.

2000–2006 Structural Funds, there was a considerable tightening of the guiding principles of community action. Among these principles, great importance has been attributed to evaluation. Structural Funds have always been an important stimulus for spreading the practice of evaluation in the European Union (Lion and Martini, 2006).

In this context, it is particularly important to see what has been done for the evaluation of the instruments for implementing rural development policy, by comparing various experiences in Italy with reference to the Rural Development Plans of regions outside Ob. 1. This research originates from the observation that, although there is undoubtedly a multitude of evaluations of the rural development planning, there does not appear to be a

similar number of studies and analyses of the «evaluation» itself (Bustelo, 2003) and in particular of analyses of complex evaluations like those of the RDPs. From recent literature, we can see a great interest in the problems of evaluation of subjects like landscape (Idda and Pulina, 2006), which are certainly more defined than the evaluation of an RDP.

Based on the contents of the mid-term evaluations (MTE) of the RDPs, the aim is to verify the level of coherence of the evaluative practice in Italy, compared to the methodological debate in recent years about programme evaluation, meta-evaluation and compared also to the EU method-

\* University of Bari, Faculty of Agriculture, Department of Agricultural Economics and Policy, Evaluation and Rural Planning.

<sup>1</sup> The paper has been accepted (and presented) at the European Evaluation Society Biennial Conference 2008, Building for the future: Evaluation in governance, development and progress, 1-3 October 2008 Lisbon, Portugal.

ological guidelines for evaluation. The main purpose of the work is to verify the methodological contents of the MTE reports of the 14 RDPs in central-northern Italy, on the basis of (meta) criteria. Through analyses and comparisons of these reports, the «quality» of evaluations is reconstructed with specific reference to methodological aspects, models and procedures adopted by evaluators.

## 2. Methodological context

### 2.1 Main approaches to evaluation

In the abundant literature on the evaluation of programmes, we find differing approaches and also numerous classifications of the proposed approaches (chronological, typological, etc.) aimed at defining the boundaries, contexts and the practices of the evaluation processes.

This contribution briefly describes three classifications (and the criteria behind them) with the aim of establishing the general reference that can be useful for analysing the theoretical scheme of the mid-term evaluation reports of RDPs.

Guba and Lincoln (1989) propose an interesting classification of evaluation approaches: 1) the technical approach; 2) the descriptive approach that qualifies activities, strengths and weaknesses with respect to the given aims (the evaluator objectively describes); 3) the approach in which the evaluator makes judgements regarding standards while maintaining the technical and descriptive functions; 4) the approach that focuses not so much on objectives, impact and decisions as on statements, doubts and points of view expressed by those involved in some way in the initiative and its evaluation. For the authors, this classification should be interpreted in historical and evolutionary terms, tracing a hypothetical pathway starting from simple evaluations and arriving at more complex ones.

The evaluation of programmes can be classified on the basis of various criteria (in addition to those previously cited), one of which could be that of the context of the evaluation. As the perspective varies, also the criteria and the procedures to achieve it clearly do so. It is therefore possible to identify four main different contexts: internal auditing, regarding evaluation of efficiency and results (usually physical); external auditing, examining specific results of policies with respect to the used resources (mainly for control and possible reduction of public expenditure); social accounting (in contrast with the limits of evaluations that only consider direct effects and financial costs), considering social and economic effects of a programme on the community; comprehension and explanation, or a detailed analysis of how programmes have achieved or not their objectives (Midmore, 1997).

Finally, in Stame's contribution (Stame, 2001), the classification of evaluative approaches is based on what is defined as a «benchmark». From this standpoint, Stame (2001) identifies three approaches: 1) experimental pragmatist; 2) pragmatist – of quality; 3) constructivist – of the social process.

For the first of these, the fundamental distinction in the evaluation is represented by the objectives declared in the programme. The evaluation consists of seeing if and how these objectives can be reached. The evaluator does not look into the processes of implementation, but waits to see the conclusion of the programme to decide whether it is worth to re-propose it. This approach was partially modified later on by including intermediate testing and monitoring of results. This, however, is limited to observing what we expected to see and is not suitable for discovering (and understanding) unexpected effects.

The approach defined as «pragmatist – of quality» is based on the comparison with a given quality standard. The evaluation becomes an exercise in comparison with a given standard. The evaluator must not be influenced by the objectives of the programme, nor he must measure how much an individual programme wants to achieve in a given situation: «the element of comparison is no longer what a single programme may want to achieve but what all similar programmes should aim at on a certain scale of merit if they want to be considered of good quality» (Stame, 2001). In this case the phase of deciding what to compare it with becomes very delicate.

Finally, for the approach defined as «constructivist – of the social process», the evaluation takes into account the results and the impact that go beyond simple achievement of aims, in order to see processes that may have been triggered. The evaluation must take into account the fact that the implementation of a programme changes in certain contexts and tries to explain why, in a given situation, it achieves certain results. This in itself is an activity that is affected by the context, which is by definition changeable. With this approach, the attention is placed on the contribution of the various parties, «and on what a programme becomes while it is being implemented, much more than on how it was designed... the important thing is that the parties are involved in a participated evaluation, that the process is followed in its various phases by the evaluators who interact with the stakeholders» (Stame, 2001). The idea of «understanding» becomes important as an aim of the evaluation.

### 2.2. Meta-evaluation: definitions, meaning, functions

The concept of meta-evaluation was introduced by Scriven in 1969, in an article called «An introduction to Meta-Evaluation» in the Educational Product Report, where meta-evaluation is defined as «any evaluation of an evaluation, evaluation system, or evaluation device». The same author went on to extend the meaning, stating that meta-evaluation means: «the evaluation of evaluations - indirectly, the evaluation of evaluators – and represents an ethical as well as a scientific obligation when the welfare of others is involved. It can and should be done in the first place by an evaluator on his or her own work; although the credibility of this is poor, the results are considerable gains in validity... [Be-

cause] the results of self-evaluation are notoriously unreliable, however, it is also desirable, wherever cost-justifiable, to use an independent evaluator for the meta-evaluation» (Scriven, 1991).

For the Joint Committee on Standards for Educational Evaluation, instead: «The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses» (Joint Committee, 1994).

Also other authors have put forward definitions of meta-evaluation, Stufflebeam and Patton being worthy of note. For Stufflebeam «Operationally, meta-evaluation is defined as the process of delineating, obtaining, and applying descriptive information and judgmental information — about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility — to guide the evaluation and/or report, its strengths and weaknesses» (Stufflebeam, 2001). For Patton too, meta-evaluation consists of «evaluating the evaluation based on the profession's standards and principles» (Patton, 1997).

Different meanings which can be given to the idea of meta-evaluation appear from the previously-mentioned definitions. A first meaning emerges from Scriven's definition. Meta-evaluation aims to examine the strengths and weaknesses of a specific evaluation activity and so is used to provide judgements about the validity and credibility of the evaluator's work. These judgments have different meanings if meta-evaluation is carried out by the same evaluator or if it is carried out by a third party. In the first case, it is an exercise in critical revision which can improve the results obtained. In the second case, it produces real value judgments to offer the «client independent evidence about the technical competence of the primary evaluator» (Scriven, 1980). With this characterization, meta-evaluation has, as its object, a single specific evaluation and is used to formulate value judgments about how the evaluation was carried out. The criteria listed in Stufflebeam's definition are then used to carry out meta-evaluation.

Carol Weiss introduces a different meaning and prefers the term «meta-analysis». She states that «Once a number of quantitative studies have been done evaluating the same kind of program, it is possible to combine their results to get better and more generalizable estimates of program effects... When the studies that are analyzed are evaluations, meta-analysis is sometimes called meta-evaluation. However, that term is properly reserved for a different concept, the evaluation of evaluations. Meta-analysis is the preferred term» (Weiss, 1998). In this case, the meaning of meta-evaluation changes. The content of meta-evaluation does not concern so much the evaluation itself, but the program or policy which is the object of the evaluation studies. The aim

consists of looking for eventual generalisations of the effects of programs (or policies). The procedures here concern statistical codification and analysis (Weiss, 1998).

Finally, for Bustelo meta-evaluation consists of evaluating the evaluation processes. «The focus is on those processes – that is, in how evaluations are done -, not on the results or findings of those evaluations. It is plausible for a meta-evaluation to include also the analysis of evaluation results, but just in the sense whether or not they are good findings (that is, if sufficient evidence was gathered, if conclusions were sounded, and interpretations, judgments and recommendations were logically drawn, and so on), and they can be useful or actually utilised for policy improvement, accountability or enlightenment. The interest of the results by themselves, that is to say in their content, is rather the focus of an evaluation synthesis and not a meta-evaluation» (Bustelo, 2002). This definition of meta-evaluation offers some interesting points compared with the others. The meaning of meta-evaluation changes yet again. In this case, it does not concern either the single evaluation or the program or policy. The object of the evaluation becomes the evaluation process. Its aim consists of «a better understanding of the evaluation function in the policy cycle and to the accumulation of knowledge in the field of evaluation» (Bustelo, 2002). In this specific area of meta-evaluation, the question which requires further investigation concerns the appropriate criteria and procedures for a meta-evaluation of different evaluations of a program or policy.

To sum up, we can identify three meanings of meta-evaluation. Firstly, the concept can be associated with the revision of a specific evaluation in order to pinpoint errors and increase the (eventual) credibility of the results. Secondly, meta-evaluation (or according to Weiss meta-analysis) consists of collecting and summarising the results of evaluations carried out on similar programs or policies or on the same program (or policy) carried out in different areas<sup>2</sup>. In this case the aim consists of collecting information which makes it possible to generalize the results obtained from a program, so as to estimate the global impact. Thirdly, meta-evaluation takes on the meaning of describing, analysing and reconstructing the overall «quality» of several evaluations. This is not done with the aim of checking or delivering judgments about the evaluations, but with the aim of investigating the specific characteristics of the evaluation function within a program or rather, in this work, to study how the evaluation process applied to the Rural Development Programs was devised and applied.

### 3. Specificity of mid term evaluation of RDPs

The total of the activities that make up the evaluation system for the development policies of the European Union has considerably evolved compared with the previous planning cycle. The system (in force for the programming period 2000-2006) is defined both in Reg. EC 1260/99 for the reform of Structural Funds and in Reg. EC 1257/99 for the

<sup>2</sup> Grajewski and Schrader (2004) carried out one of the first attempts of this kind of analysis applied to Rural Development Plans.

reform of the policy of rural development<sup>3</sup>. These regulations draw up the legal framework of reference of the evaluations of the RDPs as described below, in particular regarding the provisions on mid-term, or intermediate, evaluation of the RDPs.

The principle changes introduced by Reg. EC 1260/99 are:

- progressive improvement and increased variety of the components of the evaluation system;
- binding status;
- involvement of the operators and sharing of responsibility;
- connection between the life-cycle of the program and the multi-phase evaluation process (ex-ante, intermediate and ex-post).

From the point of view of management of the evaluation process, these changes were accompanied by the transfer of responsibility among the various operators (at a European, national and regional level). The series of subjects involved in the evaluation is therefore extended to include not only the local authorities but also the social partners and groups from society, and of course those who carry out the evaluation. In this context, MTE - that aims to give information on the level of implementation of the activities promoted by the program, on its progress and the achievement of the objectives, on the difficulties found in its implementation and possible solutions to adopt - is particularly important (Reg. EC 1260/99 Art. 42). Intermediate evaluation, in comparison with the previous periods of programming, is the phase of the evaluating process that shows the highest level of evolution because it must improve the quality and the relevance of the program. The specific «objects» of the mid term evaluations are:

- the coherence and pertinence of the program (*ex-ante* updating of the evaluation; checking of the logical framework of the program);
- effectiveness, efficiency and impact (deadweight, substitution, displacement effects);
- management, execution and monitoring procedures.

If this scenario is valid for all the programming documents of community policies, the evaluation of the Rural Development Plans, made compulsory by Reg. EC 1257/99 and by the regulations for implementation and modification<sup>4</sup>, means to guarantee the adoption of a single evaluation model and allow the comparison of evaluations of rural development in a community context. The rules regarding these evaluation activities are contained in Reg. 817/04, Art. 62-65 which deserve to be described, given their relevance. In brief, Article 62 provides that evaluations are carried out by independent experts, and that they are based mainly on a community evaluation questions and supplied with performance criteria and indicators.

<sup>3</sup> Applied and modified by following regulations: Reg. EC 1750/1999; Reg. EC 2603/1999; Reg. EC 445/2002; Reg. EC 1783/2003; Reg. EC 817/2004.

<sup>4</sup> Reg. 1260/1999 Art. 40-43; Reg. 1257/1999 Art. 48-49; Reg. 1750/1999 Art. 41-45; Reg. 445/2002 Art. 53-57; Reg. 817/2004 Art. 62-65.

Art. 64 deals with the subject of mid term evaluation. This must look at the specific aspects of the RDP under examination and at the common evaluation criteria that have relevance at a community level. Among these criteria, there are the conditions and the structure of the rural population, employment and income from agricultural and non-agricultural activity, agrarian structures, agricultural production, quality, competitiveness, forestry and environmental resources. If a common evaluation criterion is not thought to be relevant in terms of a specific RDP the reasons must be given. Furthermore, the intermediate evaluation must take into account the evaluation criteria and examine the first results obtained, their importance and their coherence with the RDP, as well as their compliance with the given objectives. The MTE must analyse the use of financial resources, the execution and monitoring of the program. Finally, it is stated that the quality of the individual evaluations is analysed by the monitoring committee - if it exists - and by the Commission, using methods recognized by the competent authorities for the management of the RDP. The results of the evaluations are made public.

In Art. 65, particularly in the third paragraph, the contents of the evaluation reports are outlined. These must show the methods applied, the implications for the quality of the data and the reports, a description of the context and the contents of the programme, the financial data, the answers to the common evaluation questions and to the questions defined on a national or regional level, with their indicators, conclusions and recommendations.

Mid-term evaluation of RDPs, therefore, takes on particular significance in understanding whether the measures implemented have been successful. It is an essential tool for rural policy.

The effort made by the European Commission to improve and, above all, standardise evaluation of RDPs can be seen in the common evaluation model based on needs, criteria and indicators. The model was defined in three working documents:

- Doc. STAR VI/8865/99 Evaluation of RDPs 2000-06. Guidelines.
- Doc. STAR VI/12004/00 Common evaluation questions and explanatory sheets.
- Doc. STAR VI/43517/02 Guidelines for the mid term evaluation.

The STAR VI/8865/99 document shows the orientations to follow in the evaluation phase, and describes the criteria and the evaluation model. The STAR VI/12004/00 document links the logical route to find the causal ties between the main variables in order to determine the results and the impact of the plan. In it, the Common Evaluation Questions (CEQ) is presented and described in general terms and in detail for every single measure. The CEQ is structured in a series of 54 questions together with assessment criteria and indicators. The evaluator applies parts of the CEQ each time they are relevant to the eligible actions and to the con-

text of a RDP; if this does not occur there must be a motivated explanation. The questions look at the effects of the programme expressed in terms of output, results and expected impact<sup>5</sup>. The answer to each question of the CEQ is based on the criteria, the indicators and on any other relevant information.

Finally, the STAR VI/43517/02 document gives clarifications regarding the intermediate evaluation that must respond to criteria of quality, utility and value. In particular this document details the specific objects of the MTE (Doc. STAR VI/43517/02, pp. 4 – 8):

checking the validity the programme strategy in relation to possible changes that have taken place in its context, in the SWOT analyses at the basis of the strategy, and of the relevance and coherence of the objectives;

analyses of the activities carried out for the start and implementation of the RDP (implementation procedures, management and organisational aspects, functioning of the monitoring system) and for the use of resources (financial and non-financial);

auditing the initial results with respect to the objectives by determining effectiveness and efficiency;

formulation of the first answers to the evaluation questions;

formulation of proposals for correction of the RDP.

#### 4. Methodological aspects of the mid-term evaluation reports

This (meta) evaluative study intends to focus on the model and on the procedures used by evaluators (and not on the results of the evaluations). We must therefore define a methodological framework of reference from which we can find the (meta) criteria for analyses (Bustelo, 2003). Regarding the model used – that is the Common Evaluation Questions – we refer to the works of Guba and Lincoln (1989), Midmore (1997) and Stame (2001). The analysis of the procedural aspects derives from the documents of the MEANS series (1999) of the European Commission and The Guide (2003) of the Tavistock Institute. These documents propose a classification of the evaluative process in four phases<sup>6</sup> (Tavistock Institute, 2003) and for every phase they show the applicable evaluative procedures. This clas-

sification has been used to identify the procedures adopted in the mid term evaluation of RDPs. The (meta) evaluative criteria have therefore been created around the evaluative model and the procedures used in the four phases of the evaluation (Diez, 2002).

#### 4.1 The evaluative model

The evaluative model proposed by the European Commission is the result of a great organisational and methodological effort and is certainly a great improvement on what was created during the period 1994-1999 for evaluating programmes. It is, however, possible to select at least two aspects worthy of comment, the first of a general character and the second regarding methodology. The inclusion of the evaluative model in the regulations has meant that RDP managers are forced into serious consideration of the evaluation of the programmes, but on the other hand it has strongly conditioned the «objects» of the evaluation, its aims, the models and even the investigative instruments.

Regarding questions of a methodological nature, it seems only right to look at how the proposed evaluative model fits into the above methodological debate. Reho, in a publication (Reho, 2006), says that «In an attempt to interpret the orientation of the Commission, while underlining the progressive move in time from an approach aimed at accountability to an approach aimed at learning, many people maintain that the EU documents give out conflicting messages. A reading of the Regulations shows the importance of accountability above all else: the Commission asks the member states to report on how the funds have been spent».

With respect to the four definitions of Guba and Lincoln (1989), the model could be placed between the second and third definition: the role of the evaluators is to describe what has or has not functioned with respect to the given objectives, but, in a certain sense, can also express judgements on the objectives themselves<sup>7</sup> (although in a limited and strongly conditioned form). Also with regard to Midmore's classification (1997), it is clear that a great effort has been made to consider the effects of the programme on communities (social accounting), at least in some of the questions of a horizontal nature in the CEQ. Finally, with respect to Stame's work (2001) it is difficult to see the real meaning of standard in the community evaluative model. Returning to Reho's thoughts (Reho, 2006), we can see how «in general the Commission tends towards an evaluation of effectiveness, comparing objectives represented on a sliding scale: the general objectives contain various intermediate objectives, each of which can be narrowed down to specific objectives, with specific operational actions... while the efficiency evaluation analyzes the relationship between input, output and results... in the background, there is the comparison of results obtained/obtainable by the programme with the initial demand, with the needs expressed by the social-economic context (external effectiveness)». It would therefore seem that one can ascribe these evalua-

<sup>5</sup> The output indicator quantifies the immediate and direct implantations, measured in physical or monetary units; the result indicator refers to the immediate but indirect effects; the impact indicator refers to the more general consequences (not immediate and indirect effects).

<sup>6</sup> The four phases are: planning and structuring of evaluation, obtaining data, analysing information, constructing evaluative judgements. The first phase consists of planning the evaluation defining the evaluative problems (the objects), the relative criteria and indicators. The second phase regards the investigation: use of primary and secondary sources. The third phase looks at the reorganisation of the information gathered in order to hypothesize first judgements and verify the hypotheses. The last phase regards the definition of the final evaluative judgements.

<sup>7</sup> Obviously in this context the term «objectives» means not only the system of RDP objectives to be evaluated (global, specific and operative), but also the sequence of questions – criteria – indicators that makes up the CEQ.

tions, with a certain caution, to what Stame (2001) defines as an «experimental positivist» approach.

In brief, we can assert that the CEQ structure uses a «traditional» approach to the evaluation of programmes, based on the quantification and measurement of implementation, results and impact. The CEQ refers to programme theory, meaning the idea that one can obtain a change, implemented through a programme, by providing resources and activities (ISFOL, 2002). The evaluative model verifies the implementation. It is substantially «goal-oriented», since it is based on an audit of the reaching of the objectives given by the programme.

## 4.2. The procedures

The main connection between reflections on the evaluative model and analysis of the procedures adopted can be found in the way in which the contents of the evaluation reports have been planned and structured (Table 1). The causal chain model (of the hypotheses) that links input to output and from which the questions, criteria and indicators proposed by the CEQ derive, comes from the «Logic Model», which is the procedure used in all the evaluation reports (except Umbria) to identify the evaluative questions of the reports.

cation of questions, criteria or indicators of the CEQ for which it would not have been possible to give an answer (with the necessary justifications required by the community document for the inability or impossibility to reply to a certain question).

If on the one hand it seems right to remember that the use of the (CEQ) model is mainly made compulsory through the community regulations, it is clear that the 14 evaluators have used little variation in their procedures. In particular, those instruments that are furthest from the outlined evaluative approach (consultation with the stakeholders to expand the objects of the evaluation or the «Formative development evaluation») have not been used. In 10 regions, however, the managing authorities and evaluators have agreed on a specific list of questions that differ from those in the CEQ, with very varied levels of extension<sup>8</sup>.

The analyses of the mid-term evaluation reports regarding the procedures adopted for gathering data and obtaining information (Table 2) generally show a sufficient diversification in the instruments used: in 10 cases, 5 to 6 procedures were used, while in just two cases very few means were used to gather information. However, the number of procedures used is not very meaningful. It is more interesting to analyse which procedures were used, and what for.

Clearly, in almost all the reports administrative data were used, i.e. data from the monitoring system. Secondary sources were mainly used for two reasons. Firstly to verify (eventual) variations in the socio-economic context, for example through the use of databases (national or regional). Secondly they were used to identify control groups in order to estimate the «net» effect of a certain measure (in particular in the RDP evaluation reports of Emilia and Friuli). For example the database of the Farm Accountancy Data Network (FADN) is used to evaluate effectiveness and efficiency through «counterfactual comparison».

In almost all the reports (except those of the RDP of Piedmont) questions are used for carrying out surveys on the recipients of the intervention. In the reports that try to evaluate the «net» effects of the intervention these surveys are carried out on a sample of the recipients. These questions are generally for gathering quantitative and qualitative data on the intervention that has been funded and impressions on the progress of the measure. Little attention is generally paid to the selection of the sample, except in a few cases (evaluative RDP reports of Marche and Emilia).

Also data-gathering based on stakeholder interviews is used in nearly all reports (except the RDPs of Abruzzo and Liguria). However, we should specify that by stakeholders we mean those responsible for the measure and the managing authorities, in other words those responsible for the administration of the programme. Interviews are rarely used to discover what other categories of stakeholders think

Table 1 – Procedures for planning and structuring the evaluation of the regional.

MTE of Regions and autonomous Provinces	Concept or issue mapping	Stakeholder consultation	Evaluability assessment	Logic models	Formative developmental evaluation
Abruzzo			X	X	
Emilia R.			X	X	
Friuli V.G.			X	X	
Lazio			X	X	
Liguria			X	X	
Lombardy			X	X	
Marche			X	X	
Piedmont				X	
Toscana			X	X	
Umbria					
Valle d'Aosta			X	X	
Veneto			X	X	
Autonomous Province of Bolzano			X	X	
Autonomous Province of Trento			X	X	

Source: our elaboration based on The Guide structure

The other procedure used (with the exception of the evaluative reports of Piedmont and Umbria) looks at the conditions of evaluability (Evaluability assessment). That is to say the analysis of the availability and usability of the monitoring system for the evaluation, and above all the identifi-

<sup>8</sup> In particular, there are specific questions in the evaluation reports of Emilia Romagna, Lazio, Toscana, Valle d'Aosta and Veneto, even though the requests of the managing authorities show a tendency to consider the evaluation of rural development as a sector, not considering the integration of the various initiatives for the development of the territory.

Table 2 – Procedures for gathering data and information in the regional evaluation reports.

MTE of Region and autonomous Provinces	Social surveys	Beneficiary surveys	Stake holder interviews	Priority evaluation	Focus groups	Case studies	Participatory approaches & methods	Use of secondary source data	Use of administrative data	Observational techniques
Abruzzo		X							X	
Emilia R.		X	X		X	X		X	X	
Friuli V.G.		X	X		X			X	X	
Lazio		X	X		X	X		X	X	
Liguria		X								
Lombardy		X	X			X		X	X	
Marche		X	X		X	X		X	X	
Piedmont			X					X	X	
Toscana		X	X			X		X	X	
Umbria			X		X			X	X	
Valle d'Aosta		X	X		X	X		X	X	
Veneto		X	X		X	X		X	X	
Autonomous Province of Bolzano		X	X		X			X	X	
Autonomous Province of Trento		X	X		X			X	X	

Source: our elaboration based on The Guide structure

about the programme. Only in the Umbrian RDP evaluation report do we find that sector experts, territorial experts and local administrators are interviewed. There is no evidence of involvement of members of other social groups (consumers, citizens etc.), nor that so-called «triangulations» have been carried out for a cross-verification of the results.

In nine evaluation reports focus groups were used usually to «enrich the quantitative data with qualitative aspects .... for example the causes that have influenced the appearance of certain phenomena, the perception of the degree of usefulness of the RDP beyond the perceived quantitative effects» (RDP evaluation report of Marche). The focus groups were usually created involving the recipients of aid and representatives of trade or business associations. In the same reports the procedure of the «panel of experts» is shown, used for the same reasons: «for the answer to some of the complex questions we use methods of «structured comparison between experts» (RDP evaluation report of Lazio and Veneto). It would seem that the focus group and the panel of experts are basically inter-changeable procedures. As it is known, the case study is a system for acquiring detailed and complete information on a certain phenomenon within its specific context, to understand complex interactions and processes. In seven evaluation reports the use of case study is explicitly mentioned, to investigate the implementation of complex projects, in particular for initiatives included in Art. 33 of Reg. EC 1257/99.

Also in this phase of the evaluation process there are interesting aspects in the procedures that do not

appear in the reports. In fact there is an absence of procedures based on structured questionnaires given to a statistically representative sample of stakeholders to obtain information for elaboration using inferential statistic instruments (a typical positivist – experimental approach procedure). The procedure defined as «Priority evaluation», based essentially on techniques of demand function estimates or identification of preferences has not been considered, not even in those areas (agro-environmental measures) where this is considered more suitable. In the same way, more qualitative investigative procedures, such as participating observation techniques, are not part of the «tool box» used by the evaluators.

To affect analysis of the information gathered and to help reach the initial evaluative judgements, in The Guide (2003) the Tavistock Institute considers input-output analysis, e-

conometric models, regression analyses, experimental and quasi-experimental approaches and Delphi survey (Table 3). GIS has been added, as indicated in the MEANS series (1999).

The first five procedures were not used in any evaluation report. No estimates were made of impacts using econometric models or analyses of the relations between variables through regressions. Only in six evaluation reports

Table 3 – Procedures for analyzing the information in regional evaluation reports.

MTE of Region Autonomous Province	Input/output analysis	Econometric models	Regression analysis	Experimental and quasi-experimental approaches	Delphi survey	GIS
Abruzzo						
Emilia R.						X
Friuli V.G.						X/Cluster
Lazio						
Liguria						
Lombardy						X
Marche						
Piedmont						
Toscana						X
Umbria						
Valle d'Aosta						X
Veneto						X
Autonomous Province of Bolzano						
Autonomous Province of Trento						

Source: our elaboration on The Guide structure

was a GIS used as an aid in analysing some aspects linked to the implementation of certain measures, in order to superimpose inferable information from monitoring on information layers available from regional cartography, or to proceed to geo-reference of the elementary data or, at least, to verify in what territorial contexts there had been a greater concentration of measures. Finally, in only one evaluation report cluster analysis was used to evaluate the presence of homogeneous groups by type of question in the analysis of agro-environmental interventions.

In the majority of the MTE reports, there is no provision for procedures to help formulate evaluative judgments (Table 4). The report from Marche is the only one that tries to apply a multi-criteria analysis to «compare the various results found in the different measures forming the Plan giving a score to the various significant aspects of the programme implementation, with the final objective of reaching a synthetic judgement on the level of efficiency reached» (RDP evaluation report of Marche). As regards the complex evaluation of the net effects of parts of the programme, we would like to underline the work carried out in the Emilia Romagna evaluation report where some indicators were calculated both on representative samples of the recipients and non-recipients («counterfactual situation»). This was an attempt to filter the sample data from exogenous influences due, for example, to the economic cycle or to external shocks (RDP evaluation report of Emilia, Appendix D).

Finally, in the evaluation reports of Lombardy and Marche there are some considerations on the implementation of the principle of equal opportunities for men and women through

the evaluation of the impact of gender on the programme. These are evaluative judgements on the adequacy of the strategy and the type of activities in the RDP with respect to the question of gender and equal opportunities.

## 5. Conclusions

In this (meta) evaluative analysis, two questions have been considered that are related to some contents of the mid-term evaluation reports of RDPs carried out in Italian regions outside Ob. 1. The first is the Community system of evaluation and its effects on the evaluative model adopted in the mid-term evaluation reports, while the second regards the evaluative procedures used by evaluators.

The evaluative system of the European Commission has undoubtedly had a great influence both on the structure of the evaluative reports and on the way in which mid term evaluation of rural development policy has been organised and managed in various regional contexts. The majority of the evaluation activities concentrate on the CEQ. In many cases, this is done to answer the questions within it (Emilia Romagna, Lazio, Toscana, Valle d'Aosta, and Veneto) and also to answer specific evaluative questions on subjects decided by administrations.

Analysis of the evaluative procedures reveals, on the whole, a limited attention to detail in the description of the procedures used and an inadequate analysis of the quality of the evaluative judgements. In the most frequent situation, in fact, the evaluators have sufficiently identified the type of instrument to use in order to structure the evaluation (although within the context of the whole approach discussed), and have carried out an adequate gathering of data and information, but have provided neither for the use of procedures for analysing the information, nor for the formulation of judgements (this aspect could also raise doubts about the soundness of conclusions).

A lack of clarity in the procedures, if not a real methodological weakness, implies the need to improve the quality of evaluations at the level of analytical methods. There appears to be a need to reflect on how to improve the means of evaluation with complex «objects» such as the evaluation of a RDP. Although there are not many specific discussions on this subject, the literature certainly has theoretical and procedural indications on the more general theme of programme evaluation. The MEANS series (1999) can be a starting point. For example, it introduces some recommendations on assembling various evaluative procedures (MEANS, 1999). The choice is

Table 4 – Procedures for formulating judgements on the regional evaluation reports.

MTE of Region	Cost-benefit analysis	Bench marking	Cost effectiveness analysis	Economic impact assessment	Gender impact assessment	Environmental impact assessment	Strategic environmental assessment	Multi-criteria analysis	Expert panels
Abruzzo									
Emilia R.				X					
Friuli V.G.									
Lazio									
Liguria									
Lombardy					X				
Marche					X			X	
Piedmont									
Toscana									
Umbria									
Valle d'Aosta									
Veneto									
Autonomous Province of Bolzano									
Autonomous Province of Trento									

Source: our elaboration on The Guide structure



guided by the answers to some questions, the main one being: «Are the procedures appropriate for the phases in which they are being used?». It is clearly not our intention to maintain a position in which it is not clearly underlined that «it is necessary to be able to use all the range of approaches, to be able to combine methods and instruments» (Stame, 2001), particularly in evaluation of programmes that aim to implement a policy like the one (multidimensional) of rural development; all this, being fully aware that the quality of an evaluation does not exclusively depend on the methods or on procedures used but on the way these are used to produce useful and useable information (judgements).

In this (meta) evaluative analysis, the main definitions, meaning and functions of meta-evaluation have been introduced. Meta-evaluation has been intended as a tool to investigate in the evaluative process. Clearly, another meaning of the meta-evaluation is to verify the quality of the evaluation and of the evaluators work. But, whatever the meaning of meta-evaluation we may consider, the practice of meta-evaluation is not diffuse. Stufflebeam (2001) says that: «Meta-evaluation is as important to the evaluation field as auditing is to the accounting field. Society would be seriously at risk if it depended only on accountants for its financial information, without acquiring the scrutiny of independent auditors. Likewise, parents, students, educators, government leaders, business persons, and consumers, in general, are at risk to the extent they cannot trust evaluation findings». Structural Funds should undertake the role of stimulus for spreading the practice of meta-evaluation.

## References

Bustelo M., 2002. Meta-evaluation as a tool for the improvement and development of the evaluation function in public administrations. Paper presented at, 5th European Evaluation Society biennial conference: Learning, Theory and Evidence, 10-12 October 2002, Seville.

Bustelo M., 2003. Evaluation of Gender Mainstreaming. Ideas from a Meta-evaluation Study, *Evaluation* 9(4), 383-403.

Diez M.A., 2002. Evaluating New Regional Policies. Reviewing the Theory and Practice, *Evaluation* 8(3), 285-305.

European Commission, 1999. MEANS Collection. Evaluation of socio-economic programmes - 6 volumes. Luxembourg. Office for Official Publications of the European Communities.

European Commission, 1999. Evaluation of RDPs 2000-06. Guidelines. STAR VI/8865/99.

European Commission, 2000. Common evaluation questions and explanatory sheets. STAR VI/12004/00.

European Commission, 2002. Guidelines for the mid term evaluation. STAR VI/43517/02.

Grajewski R. and Schrader H., 2004. Design, results and lessons from program evaluation: evidence from the mid term evaluation of six rural development programmes in Germany, Selected Paper at the 87th Seminar of the EAAE 'Assessing rural development of the CAP', 21-23 April 2004, University of Agricultural Sciences, Vienna (Austria).

Guba E.G. and Lincoln Y.S., 1989. Fourth generation evaluation, Sage Publications.

Idda L. and Pulina P., 2006. Paesaggio e sviluppo rurale in Sardegna, FrancoAngeli.

ISFOL, 2002. Orientamenti metodologici per la valutazione del processo di implementazione di programma, Collana Metodologie per la valutazione di programma, volume 2, UE-MLPS-Fse, Roma.

Joint Committee on Standards for Educational Evaluation, 1994. The Program Evaluation Standards, 2nd ed. Thousand Oaks, Sage.

Lion C. and Martini P., 2006. The evaluation of a Complex Social Program: Lessons learned from the experience of the European Social Fund, *Evaluation and Program Planning*, 29, 1-9.

Midmore P., 1997. Rural Policy Reform and Local Development Programmes: Appropriate Evaluation Procedures, *Journal of Agricultural Economics*, 49(3), 409-426.

Patton M.Q., 1997. Utilization-Focused Evaluation, Thousand Oaks, Sage.

Reho M., 2006. Le misure per la tutela e la valorizzazione del paesaggio introdotte dalla nuova PAC. Valutazioni di efficacia in relazione ai fattori di contesto e alle modalità di gestione, In Marangon F., (a cura di), *Gli interventi paesaggistico-ambientali nelle politiche di sviluppo rurale*, FrancoAngeli.

Scriven M., 1969. An Introduction to Meta-Evaluation, *Educational Product Report* 2(5), 36-8.

Scriven M., 1980. *Evaluation Thesaurus*. Second Edition, Edgepress Inverness California.

Scriven M., 1991. *Evaluation Thesaurus*. Fourth Edition, Newbury Park, Sage.

Stame N., 2001. Tre approcci principali alla valutazione. In Palumbo M, (a cura di), *Il processo di valutazione. Decidere programmare valutare*, FrancoAngeli.

Stufflebeam D.L., 2001. The Metaevaluation Imperative, *American Journal of Evaluation* 22(2), 183-209.

Tavistock Institute, 2003. *The GUIDE – The Evaluation of Socio-Economic Development*. <http://www.evaled.info>.

Weiss C.H., 1998. *Evaluation*. Second Edition, Upper Saddle River (NJ), Prentice Hall.