# METAEVALUATION OF THE FIRST CYCLE OF EVALUATIONS OF THE QREN (NSRF) AND THE OPERATIONAL PROGRAMMES (2007-2013)

## EXECUTIVE SUMMARY
### DECEMBER 2012

| | |
|---|---|
| **TITLE:** | EVALUATION REPORT OF THE QUALITY AND USEFULNESS OF CONSTANT EVALUATION PROCESSES OF THE GENERAL EVALUATION PLAN FOR THE QREN (NSRF) AND THE OPs 2007-2013 [METAEVALUATION] |
| **SUBJECT** | CONSTANT EVALUATION PROCESSES OF THE GENERAL EVALUATION PLAN FOR THE QREN (NSRF) AND THE OPs 2007-2013 <br><br> INCLUDING THE 15 (FIFTEEN) EVALUATIONS OF THE QREN (NSRF) AND THE OPs ALREADY CONCLUDED AT THE TIME OF THE TENDER [11/07/2011] |
| **GRANTING ENTITY** | OBSERVATÓRIO DO QREN |
| **ENTITY RESPONSIBLE FOR EVALUATION** | INSTITUTO DE ESTUDIOS DE ECONOMÍA, EVALUACIÓN Y EMPLEO, S.L. [*I3E*] <br> www.ie3e.es |
| **EVALUATION TEAM** | **JOSE LUÍS OSUNA LLANEZA.** UNIVERSITY OF SEVILLE**. [SCIENTIFIC DIRECTOR]** <br><br> **JUAN MURCIANO ROSADO.** I3E - UNIVERSITY OF SEVILLE**. [COORDINADOR]** <br><br> **ESTER GARCÍA SÁNCHEZ.** IPOLGOB – CARLOS III UNIVERSITY OF MADRID <br><br> **CARLOS BUENO SUÁREZ.** UNIVERSITY OF SEVILLE <br><br> **FÉLIX RODRÍGUEZ VERANO.** I3E <br><br> **JUAN FRANCISCO CASERO CEPEDA.** I3E <br><br> **BELÉN SOTILLOS GONZÁLEZ.** I3E COLLABORATOR <br><br> **JOSÉ MANUEL BETANZOS MARTÍN.** I3E COLLABORATOR |

## ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| MA | Managing Authority responsible for evaluation launching and implementation |
| WORKING DOCUMENT No 5 [WD5] | WORKING DOCUMENT No 5 "INDICATIVE GUIDELINES ON EVALUATION: EVALUATION DURING THE PROGRAMMING PERIOD. GUIDANCE DOCUMENT OF THE EUROPEAN COMMISSION CONTAINING GUIDELINES ON OBJECTIVES, ACCOUNTABILITY AND KEY ASPECTS OF EVALUATIONS TO BE CONSIDERED IN THE CURRENT PROGRAMMING PERIOD IN ORDER TO FACILITATE TO OP MANAGERS COMPLIANCE WITH THE REQUIREMENTS LAID DOWN IN ARTICLE 47. |
| ERDF | European Regional Development Fund |
| CF | Cohesion Fund |
| SG | Steering Group |
| IFDR | Instituto Financeiro para o Desenvolvimento Regional, IP |
| IGFSE | Instituto de Gestão do Fundo Social Europeu, IP |
| General Guidelines | General Guidelines for the implementation of Operational Plans |
| EP | Specific evaluation plan for an OP or for the NSRF |
| GEP | General Evaluation Plan for the NSRF and the OPs |
| OPS | Operational Programmes |
| NSFR | National Strategic Reference Framework (Quadro de Referencia Estratégico Nacional e dos Programas Operacionais) |
| REGULATION (EC) No 1083/2006 | COUNCIL REGULATION (EC) No 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999. Conselho, de 11 de julho de 2006, que estabelece as disposições gerais relativas ao Fundo Europeu de Desenvolvimento Regional, ao Fundo Social Europeu e ao Fundo de Coesão e que derroga o Regulamento (CE) Nº 1260/1999. Corrigendum07-06-2007. |
| ToR | Terms of Reference of the evaluation (Specifications, Tender Document, Call for Tenders,…) |

## CONTEXT OF THE METAEVALUATON

The European regional policy pursues social and territorial cohesion by a reduction of the development disparities between regions. To this end the Structural Funds and the Cohesion Fund have been established as financing instruments and the "Quadro de Referencia Estratégico Nacional (National Strategic Reference Framework, hereinafter, NSRF) and the Operational Programmes as instruments for planning, management and implementation in each Member State. From its origins it has shown a marked concern to know accurately the destination of the funds used and hold users accountable for the achievements made. Thus, evaluation is a requirement of community funding that has promoted evaluation culture in the Member States.

Indeed Council Regulation (EC) No 1083/2006 establishes evaluation objectives, when such evaluations shall be implemented and details the responsibilities of Member States, among them drafting of an Evaluation Plan, ex-ante evaluation of OPs and their evaluation during the programming period. The latter results in a Metaevaluation for which this Summary is presented.

## OBJECTIVES, OPPORTUNITY AND METHODOLOGICAL APPROACH

The **general objective** of the final Metaevaluation Report of the first cycle of operationalisation[1] of Operational Programme and 2007-2013 NSRF evaluations[2] is to determine the quality and usefulness of these evaluations so that goodness and merit of the Evaluation system in which they are framed can be assessed and recommendations can be made to improve them for the new EU programming period.

The **subject matter of the analysis** is evaluation processes, as well as the whole range of contextual factors that impact on them, especially the organization and operation of the system in which they take place.

In line with the above, its **specific objectives** are to **analyse and assess both in the evaluations and in the system**:

- **Clarity and coherence** of the design and assignment of evaluations in order to optimise processes[3].

- **Transparency** in the procurement process (publicity of procedures and independence of evaluation teams).

- **Sufficiency** of human, financial and time resources allocated to processes.

- **Accuracy** to provide valid and reliable information with relevant and robust methods and coherent conclusions and recommendations.

- Ability to ensure stakeholder **involvement** integrating their perspectives and opinions and triggering individual and organizational **learning** processes.

- **Usefulness** to provide responses to stakeholders' information needs.

---

[1] Being this the first evaluations carried out after launching the Operational Programmes, their purpose is to determine the contribution of the OP management system and its operational mechanisms (candidate selection criteria, process handbooks, etc) to the achievement of the strategic priorities –in which priority axes are extremely relevant- and an appropriate articulation of operational and strategic elements.

[2] The full name of each of the evaluations, as well as their code and the short name used are as follows: AVALIAÇÕES GLOBAIS: (1) Avaliação Global da Implementação do QREN 2007-2013 [QREN 2007-2013]; PROGRAMAS OPERACIONAIS TEMÁTICOS: (2) Avaliação da Operacionalização do POPH 1.4 e 9.1.4- Cursos de Especialização Tecnológica [POPH 1.4-9.1.4]; (3) Avaliação da Operacionalização do [POPH 2.3, 8.2.3 e 9.2.3- Formações Modulares Certificadas [POPH 2.3-8.2.3-9.2.3]; (4) Avaliação da Operacionalização do POPH 3.2, 8.3.2 e 9.3.2 -Formaço para a Inovação e Gestão [POPH 3.2-8.3.2-9.3.2]; (5) Avaliação da Operacionalização Inicial dos Sistemas de Incentivos [SISTEMA DE INCENTIVOS]; (6) Avaliação da Operacionalização do POVT [POVT]. PROGRAMAS OPERACIONAIS REGIONAIS DO CONTINENTE: (7) Avaliação da Operacionalização do ON2 - "O Novo Norte" [O NOVO NORTE]; (8) Avaliação da Operacionalização do Mais Centro [MAIS CENTRO]; (9) Avaliação da Operacionalização do POR Lisboa [POR LISBOA]; (10) Avaliação da Operacionalização do Inalentejo [INALENTEJO]. PROGRAMAS OPERACIONAIS DAS REGIÕES AUTÓNOMAS: (11) Avaliação da Operacionalização do PO PROCONVERGENCIA [PROCONVERGENCIA]; (12) Avaliação da Operacionalização do PROEMPREGO [PROEMPREGO]; (13) Avaliação da Operacionalização do Programa Intervir+ [INTERVIR]; (14) Avaliação da Operacionalização do Programa Rumos [RUMOS]. AVALIAÇÕES ESPECÍFICAS: (15) Avaliação do Modelo de Elaboração dos Regulamentos Específicos dos Programas Operacionais FEDER e Fundo de Coesão 2007 - 2013 [FEDER e FC].

[3] The focus lies one the main documents of the regulatory and procedural framework: Guidelines of the Evaluation Network, Global Evaluation Plan, individual Evaluation Plans, Terms of Reference and Evaluation Proposals.

If the **timeliness** of any *metaevaluation* lies in its ability to determine the quality of evaluations, promoting learning among evaluators and managers, improving professional quality and strengthening accountability to citizens, in this case it is further reinforced because it is a key input to guide and optimize evaluation tasks in the 2014-2020 programming period, which will face new challenges also in the field of evaluation.

Despite of the positive aspects of metaevaluation, its practical implementation is still scarce and quite limited to the academia, especially in countries with less evaluation tradition and culture and still incipient institutionalisation processes. This provides this work with a marked **innovative nature**, moreover since it is considered as a tool to improve management, effectiveness and impacts of Structural Funds in Portugal.

In order to tackle these challenges this metaevaluation[4] combines the analytical, quantitative and objective nature of evaluation based on criteria with the use of qualitative methods aimed at developing an interpretative component based on subjectivity, unique contexts and individual experience (Stake, 2006).

For this reason it was decided to employ **triangulation** as the research strategy, combining different techniques to enhance the validity of its results and capture a more comprehensive image of how evaluation is being structures in the NSRF. Work has been structured into **three sequential stages**.

During the Preliminary Stage a **documental analysis** was carried out, addressing the regulatory framework, general guidelines and the documental support of each evaluation. A total 105 documents were analysed to gather primary, objective and comparable information which was entered into a database. The results have allowed for a context-adapted design of the tools to be applied in the second State (information gathering, treatment and analysis).

In the second stage a series of semi-structured interviews were conducted to evaluation managers (16) and evaluators (4), which contributed key qualitative elements. They were face-to-face interviews (except for one videoconference) and they were recorded in order to allow for their transcription and the analysis of their content. Furthermore, two strategic participation workshops were organised with evaluation managers, one focused on diagnosis of the evaluation system and the other on conceptualisation of the key factors for high quality evaluation in the NSRF context. Participation of all the attendants was encouraged and group production was optimised by using cards filled out by all participants that were shown on slides in order to get feed-back and enhance collective reflection.

Finally, a survey was conducted to the evaluation manager and evaluator population involved in the 15 evaluation. The self-administered online questionnaire included 15 items on the factors determining a high quality evaluation, and asked them whether they agreed or disagreed in a 0 to 10 scale for each evaluation in which the survey respondent had been involved. The sample was made up by 30 representatives of the 15 evaluations (15 managers and 15 evaluators). 33 completed questionnaires were returned accounting for 14 of the 15 evaluations (93%) and half of them were assessed both by managers and evaluators. 30% of the completed questionnaires were returned by evaluators and 70% by managers. Finally, we should point out that some of the limitations were the information deficits in some of the evaluations, the difficulty to cover the whole range of stakeholder profiles and that the field work took place at a time in which OPs required a large amount of management work.

The Final Stage included the analysis and interpretation of the information gathered. The main findings were used to draw the resulting conclusions and recommendations.

---

[4] The theoretical framework of this Metaevaluation incorporates proposals enjoying the highest international reputation: Schwandt and Halpern (1988); Daniel Stufflebeam (1999); Stufflebeam, Goodyear, Marquart and Johnson (2005); Üusikylä and Virtanen (2000); House and Howe (2000); y Stake (2006).

## METAEVALUATION CONCLUSIONS AND RECOMMENDATION

This main section contains an assessment of the main results from the meta-evaluation. It includes the main conclusions (C) and recommendations (R) classified according to each of the three evaluation stages and for each of the evaluation questions and criteria detailed in the Terms of Reference.

**EVALUATION STAGE: (I) CONCEPTUALISATION AND ASSIGNMENT**

**1.1 CLARITY AND COHERENCE:** Was the evaluation conceptualisation process sufficiently clear and consistent as to maximise evaluation processes?

C1 The documents on evaluation conceptualisation and assignment are clear in their wording and content and are highly coherent with their reference framework; however, there are some areas that may be improved to maximise evaluation processes.

- **General Guidelines and the GEP** are highly coherent with the reference framework established by Regulation (EC) No 1083/2006, Working Document No 5 (EC) and EVALSED Guidelines, as it has been confirmed by a documental analysis of the sections on objectives, principles, type of evaluation, accountability, budget, methodology and dissemination. Moreover, such clarity and coherence has been unanimously supported by all evaluation managers, who highlighted their usefulness for the drafting of Specifications at interviews and workshops.

- **Evaluation Plans** are coherent with the General Guidelines and the GEP. The general contents that have been best defined are those on objectives, questions and support structures and the interactive process used in their drafting is considered very appropriate and should therefore be continued in the future. The main deficit that has been detected is the limited specificity and level of detail in some of their contents beyond the GEP while others that would contribute to make evaluation more specific are absent. Field work supports the good quality of the EPs and key stakeholders have got a positive perception of them, which is even better among managers as evidenced in the survey. They made special emphasis on the collaborative and interactive EP and GEP design process, which has enhanced networking and skill building among them.

- **Terms of Reference (ToR)** are coherent with the Evaluation Plans, the General Guidelines and the GEP. A majority of the interviewed subjects agree that the Terms of Reference are useful guidelines to draft the proposals. The level of agreement is lower among evaluators, who are the target group. With regards to stakeholder participation in this process as a way of including their information needs and developing truly collaborative processes, 35% of the evaluators that have been interviewed declared that it was insufficient, whereas 75% of the managers answered that they fully or very much agreed on their sufficiency.

- **Evaluation Proposals** are coherent with the Terms of Reference regulating them. The documental analysis confirms such coherence for part of the evaluations included in the analysis, since they meet both formal and content requirements, especially those established in the General Guidelines for the two public procurement procedures under study. Coherence and clarity of the proposals has been supported by interviewed managers who, although they declared in the survey and the interviews that they felt that there was room for improvement, especially in terms of methodological innovation and that they met the required quality standards and were useful guidelines in the implementation of evaluations.

### RECOMMENDATIONS

R1 General Guidelines should explicitly include PA design and not only their implementation ["General Guidelines for PA design and implementation"].

R2 OP target groups and other stakeholders should be included in the design of the evaluation plan or, at least they should be asked about their information needs on evaluation. Furthermore, their participation in this initial stage of the Steering Groups should be enhanced.

R3 A document describing the Evaluation Policy of the QREN (Portuguese National Strategic Reference Framework - NSRF) context should be drafted.

**1.2 TRANSPARENCY:** Was the evaluation team selection process accurate and transparent?

C2 The evaluation team selection process was accurate and transparent as a result of the public procurement procedures used and in all cases evaluations teams external to the implementation, certification and auditing bodies of the NSRF context. However, the predominance of direct assignments (73%) over public tenders (27%) as the optimal choice has had a negative impact on the transparency score.

**Publicity in procurement.** The preference for Direct Assignment was associated to the fact that managers focused on the baseline price of this procedure and on the results obtained as compared to those achieved by means of Public Tenders. In this way, they have adopted a utilitarian concept of the procurement process –in terms of cost-benefit or according to efficiency-effectiveness criteria- which took precedence over one based on moral and ethical aspects according to the legitimacy criteria of the evaluation process which are intended to grant greater transparency.

Notwithstanding the above, we consider that the best option for assigning an evaluation is through Public Tender because it is the procedure providing greater control and publicity in public procurement and as a result greater transparency. Furthermore, if the aim is to assess the merits or value of a public program, such process adds legitimacy to the winner, since although the intention of seeking at least three teams in Direct Assignment is aimed at guaranteeing quality alternatives, it turns out to be a discretionary decision. Such decision may be based on merit and value principles basic to any public procurement process; however, it does not ensure equal opportunities, a likewise essential element.

**Independence** of evaluation teams from management, certification and audit authorities has been taken into account in all evaluations since all winning teams were external to them. In the actual area of operationalisation evaluations it has resulted in resorting only to external evaluation. Despite being external to such authorities, during the interviews it was mentioned that in some cases there were links to policymakers, because they had previously held political positions, or related to the programmes, or were involved in their design. However, no references were found in the regulatory framework preventing the provision of services in such circumstances and there is no specific reference to then as a conflict of interests regarding evaluation.

**Participation** of stakeholders during the evaluation assignment and contracting process allowed them, according to the interviewed, to incorporate their information needs into the process; the most frequent score was 8 over 10 with a 6.58 average. The score was higher among managers than among evaluators. Notwithstanding, in the interviews it was suggested, especially by evaluators, that evaluations reflect mainly issues of interest for OP managers. They considered that although they should take precedence, being operationalisation the subject matter of evaluation, issues of interest for others should not be ignored.

## RECOMMENDATIONS

**R4** Measures aimed at evaluation publicity and independence should be enhanced

- R.4.1 Setting up thematic working groups in the GEP public presentation day with the participation of entities and university departments specialised in OP target topics and areas and in evaluation.
- R.4.2 Establishment of a Register/repository of NSRF Evaluators in order to allow Managing Authorities to detect potential specialised teams.
- R.4.3 Completing a legal feasibility study on the incorporation of specific situations as conflict of interests into the policy framework.
- R.4.4 Exploring during the new programming period, pilot experiences of internal/combined evaluation for operationalisation evaluations, but not in interim or outcome evaluations, which are more appropriate for external evaluation.
- R.4.5 Maintaining the general guideline on preference of public tenders for the evaluation procurement process, being this the most appropriate procedure inasmuch as it provides a more open market and ensures equal opportunities for potential evaluation teams.
- R4.6 All the information related to the evaluation procurement process should be posted on the website (of the programme under evaluations and of the Observatory), regardless of whether it is published on other in Public Administration procurement platforms.

**R.5** Promotion of the evaluation culture should be included as work line in order to strengthen the evaluation market.

**1.3 SUFFICIENCY: Does the Evaluation System have enough resources for appropriate launching and implementation of evaluation processes?**

**C3** Resource management is causing some tensions in the Evaluation System that may compromise its quality and medium/long-term feasibility resulting from:

- poor evaluation scheduling –it underestimates where it does not ignores the tasks of evaluators and the obligations of managers and Steering Groups in the process- leading to very significant accumulated delays;
- budgeting is determined by procurement procedures rather than by evaluation requirements, and there is a poor adjustment of expectations and questions-; and
- absence of an organically and functionally specialised professional profile in evaluation management and Management Authorities, together with the scarce supply of specialised teams in Portugal.

**Timing**. Documental analysis has confirmed that evaluations have accumulated very important delays as compared to planned deadlines: 1 quarter in the Launching stage; 2.6 quarters in the Implementation stage; 8.5 months in the submission of reports (over the deadlines established in the GEP) and 10.6 (over the deadlines of the ToR), taking into account that half of the evaluations established in the ToR shorter terms that planned (GEP).

The reasons justifying such delays are connected to the tasks of evaluation managers and evaluation teams, according to the information gathered during field work. The survey has analysed these perceptions and managers are who give the worst scores to the expected duration of the evaluations.

**Budgeting**. Documental analysis has confirmed the reduction of the budget; not only between programming periods but between the disclosure dates of the GEP/EP and the disclosure dates of the ToR. According to the initial GEP, the planned average budget was €75,000, whereas according to the updated version of the GEP, the effective average budget was €58,544 (22% less). When evaluations of the NSRF, Bonus System and FEDERFUNCO are excluded, this actual average budget is further reduced to €38,894 (48% less). Likewise, interviewed stakeholders agree in pointing out the increased demands of evaluations as compared to the previous programming period (2000-2006). Documental analysis has confirmed that the requirements (questions) had not been reduced between the GEP and the publication of the ToR. Interviews have also evidenced the potential threat posed by the fact that those specialised teams may not be viable or that the quality of the field work performed may be lower. The survey has allowed for an analysis of such perceptions: The question of whether the effective budget allowed for an appropriate performance of evaluations got an average score of 6.64, reporting one of the largest discrepancies between managers and evaluators (7.78 and 4).

**Human Resources**. Field work has revealed that on the evaluation demand side, in most of Management Authorities there are no specialised teams in evaluation, not even a person devoted exclusively to this task, which causes tensions and may relegate evaluation processes to a second level. On the supply side, it has been confirmed that a small number of entities have performed most evaluations, a consequence of the small size of the market in Portugal, something on which most of the interviewed agreed. Furthermore, evaluators considered that their evaluation skills were optimal while managers evidenced the difficulties they face in managing them.

The consequence of this scenario is that, on the one hand, overspecialisation of these evaluation teams has leas them to take part at these evaluation "at any cost" and on the other, Managing Authorities, given the impact of the crisis on reprogramming and the reduction of available resources, have established baseline prices which do not consider the actual requirements. For the new scheduling period these tensions could grow and pose an important threat for the sustainability of the system itself.

## RECOMMENDATIONS

**R6** Optimal time required for evaluations should be determined as to match the demanding requirements and expectations as well as the available budget. A time schedule should be established for a "standard" evaluation process with a holistic and pluralist perspective.

**STAGE OF EVALUATIONS: (II) IMPLEMENTATION**

**ACCURACY:** Did evaluations reveal technically accurate and trustworthy information about the aspects determining the merit and value of established objectives?

**2.1 THEORETICAL METHODOLOGICAL FRAMEWORK**

**C4** The scarce importance attached to theoretical methodological aspects is one of the main weaknesses of the evaluations under study.

This is evident not only by the weakness of the methodologies used -according to managers- but also by the fact that the final reports did not include –or when included it was clearly insufficient or focused only on data collection- a section explaining which is the evaluation approach used or what is the methodological design of the evaluation.

The reason for this is that at the end of the day an evaluation is a *scientific research process*. Its specific features as compared to other research tools and types –it has public policy actions as its subject matter, the evidence obtained is used to judge the value of such actions and to introduce timely improvements- though absolutely essential, do not diminish the validity of the argument but rather the opposite. If the purpose is to improve public policies based on sound (scientific) evidence provided by the evaluation, this should necessarily rely on solid foundations and to this end it is necessary to have a consistent and well-articulated theoretical system in place.

With regards to the reasons that could explain such deficiency, data are not indisputable and just allow for venturing a hypothesis, being the most plausible the one that connects them to the poor development of evaluation in the country and as a result training deficits on the matter. This could explain why evaluations have had a rather limited assessment nature in some occasions and that they resembled consultancy work on public policy actions.

## RECOMMENDATIONS

**R7** Investment (not only financial) in training could constitute a key factor to encourage the development of evaluation.

## 2.2 OBJECTIVES

**C5** In general terms, evaluation objectives and questions are perceived as clear and they allow for an appropriate implementation of evaluations.

Nevertheless, evaluators and managers believe that there is some room for improvement, especially with regards to the questions, as confirmed by the data obtained from the survey –which shows the evaluators are more critical about this topic- and from the workshops –where the managers identified the ambiguity and lack of focus of the questions as one of the main weaknesses of the evaluation processes.

### RECOMMENDATIONS

**R8** Continuous attention shall be paid to the determination of evaluation objectives, as their clarity and precision are considered to be satisfactory for the implementation of evaluation processes.

## 2.3 CRITERIA

**C6** The presence of value criteria is very scarce in the evaluation processes that have been analysed

An explicit used of value criteria is essential because it is *the criteria –and not the questions* in which they are detailed- *that are the references or patterns based on which a public policy action is judged as a success or a failure*. In short, "an evaluation is not an evaluation as such if it does not conclude with a value attribution." Making a specific statement on the criteria guiding the process allows for asking the right questions, provides a evaluations with accuracy and a systematic nature and leads to sound conclusions: "Without some value criteria clearly stated from the beginning, the final report of the evaluation will barely offer any conclusion." This also constitutes a guarantee of transparency – especially if they are disclosed to the public- because, where necessary, it allows for meta-evaluation processes and it allows audiences and stakeholders to know which are the "values" determining the success or failure of public policies.

### RECOMMENDATIONS

**R9** A minimum set of criteria providing the backbone for Operative Programme evaluation and judgment processes should be established.

## 2.4 EVALUATION QUESTIONS

**C7** Evaluation questions are one of the weakest elements in the evaluation processes that have been analysed because of their quantity more than their quality although their focus could also be improved.

It should be noted that the importance attributed to evaluation questions in the Evaluation System is one of its main strengths. Having said this, the evaluation proposal is perceived, mainly in terms of the amount of topics covered, as too ambitious, especially considering their budget allocation and estimated implementation time, although it is considered as a positive element the fact that the Terms of Reference specify such issues because it simplifies the evaluator's work.

Data from the field work do not allow us to state firmly that the wording of the questions is imprecise and ambiguous, and that this has negatively conditioned evaluations. The information provided by the survey does not coincide with that coming from the workshops, which shows a more negative opinion about this matter. However, this topic was mentioned several times during the first participatory workshop. There is no doubt that the questions do mostly reflect the managers' requests for information, while they barely deal with the needs that, in this sense, could have other stakeholders. Thus, the potential use of evaluations is reduced because it is impossible to learn about public policies under which the programmes are implemented, which may be fundamental for the coordination and the internal coherence of the same policy.

### RECOMMENDATIONS

**R10** An evaluation question priority setting process should be established to structure the design of the GEP-EP-Terms of Reference.

- R10.1 Including in the General Guidelines a specific section: "'Evaluation questions: importance, wording and priority setting.'"
- R10.2 Including the wording and priority setting of evaluation questions in EPs.
- R10.3 Developing evaluation managers' skills to undertake formulation processes and evaluation questions.

## 2.5 SOURCES, TECHNIQUES AND FIELD WORK

**C8** OP information systems have received little attention in the Terms of Reference, evaluation proposals and reports, despite of the fact that usual discourses have evidenced the importance attached to them as a condition necessary to develop optimal evaluation processes, affecting the techniques which have to be used and the proposal nature the same evaluations.

If the discourses agreed on the qualitative step forward of these Systems during the current programming period, there is a gap between the opinion of managers and evaluators regarding the room for improvement in them as conditioning elements of any evaluation process. Whereas managers are satisfied with their current content and the extensive possibilities to make use of them in evaluations, evaluators point out that this it is true in many cases and circumscribed to certain areas, especially the existing candidatures; even though they point out at other bottlenecks connected to series discontinuity, lack of updating, disaggregation level or homogenous treatment by all the OPs, together with other technical and IT difficulties. These perceptions are coherent with many of the recommendations made in the evaluations themselves, which aimed precisely at an improvement and standardisation of such Systems.

Although it is not the objective of this work, the interpretation of these findings is that Systems focus mainly on meeting the information requests of OPs regarding the policy and procedural framework and its auditing and accountability processes, rather than those of the evaluation processes themselves. This was supported by the statements made by several interviewed managers and evaluators.

**C9** The information gathering and analysis techniques considered in the evaluations are appropriate for the evaluation objectives and questions. They offer multi-methodological approaches with a positive balance of qualitative and quantitative techniques and focus on involving and consulting key stakeholders to the programs.

The information gathered allows us to conclude that it is relevant in terms of type and amount for the operationalisation evaluation objectives and questions and coherent with the requirements established in the Terms of Reference. This is so aside from the fact that since a range of social research techniques are used in the evaluation it is always possible to hypothesise other potential ones and that such process may be considered to be 'innovative'.

**C10** In general terms techniques have got an internal coherence in their conceptualisation although some shortcomings were detected in their application, especially in the case of the surveys –low response rates of answers or excessive margin of error- but also in the group dynamics –panels and focus groups, due to the small number of participants-; notwithstanding, this did not have a substantially negative and overall impact on their robustness and subsequent credibility were due to the variety of techniques applied.

Although interviews did not show such blunt differences, the survey has shown that there are different perceptions regarding the accuracy of the evaluations conducted, being this one of the main conflict areas between managers and evaluators. Evaluators strongly or fully agreed that the methodology and techniques were appropriate, while managers reported a medium level of agreement, as 30% of them strongly disagreed. According to the interviews, this critical position is linked to a unanimous perception about 'lack of innovation' in this matter and that they are considered as routine work in evaluations.

This perception of managers combines with the lack of an accurate delimitation across the System as a whole of the theoretical approaches of evaluation, methodology and techniques and the subsequent consequent lack of attention paid to the available guidelines and tools. In part, it is true that many others could have been considered; although it is also true that technique selection is a second step after methodological design. This is an essential area that is not explored when it comes to **innovation** in the NSRF context; not only by managers -associating innovation to techniques- but also by evaluation teams. Thus, we consider that the niches for innovation can be found more in the theoretical-methodological framework and the exploration of evaluation models and approaches than in actual techniques applied. With regards to the latter, it is not surprising that proposals are very similar considering that evaluation contexts and subjects are very similar and once they have proven to be valid they are used unless otherwise required. Although this is not an obstacle for exploring other alternatives, it cannot be considered as a factor limiting their pertinence and accuracy.

## RECOMMENDATIONS

**R11** Requirements of methodology and techniques should be optimized in the documents detailing the conceptualisation and assignment of evaluations in the System.

- **R11.1** Incorporating "Treatment of the triangulation strategy"' into the evaluation proposals and reports in an explicit manner, which requires the use of multi-method approaches.

- **R11.2** Incorporating a justification of the "'Selection of techniques used" into the evaluation proposals (and final report).

- **R11.3** Incorporating a specific statement on the "Limitations of evaluation" into the evaluations (Final report).

**R12** Steering work should be reinforced during the implementation of evaluations and the organisation of field work, and managers' experience in this area should be systematized.

**R13** A work line aimed at Improving Information Systems in Operative Programs should be established for evaluation purposes.

## 2.6 EVALUATION CONCLUSIONS

**C11** The evaluation reports that have been analysed provide coherent conclusions, supported by evaluation findings. Their wording is clear although they tend to use an excessively technical and encoded language is used. Classification and structuring criteria are heterogeneous criteria which does not encourage collective learning from completed exercises within the Network.

Document analysis evidences the aforementioned coherence and clarity. However, during the survey different perceptions were detected among managers and evaluators, evidencing there is a wide margin of improvement regarding conclusions. This discrepancy was not evident during interviews and workshops as debates focused always on recommendations and there was little, if any, on conclusions. It should also be taken into account that, in average, only 17% of the recommendations was rejected, assuming that those conclusions based on evaluators' prejudices not supported by data and evaluation findings would have been rejected by managers.

### RECOMMENDATIONS

**R14** Any potential measure in the System aimed at providing strong foundations for evaluation conclusions should be developed.

- **R14.1** General Guidelines should include basic methodological standards in order to make basic formal and content requirements more explicit when it comes to drawing conclusions; this would make it easier to in transfer them to the Terms of Reference.
- **R14.2** The necessary links and structural relations between evaluation questions and conclusions should be reinforced, making use of all means to ensure them (grouping, numbered presentation and, whenever possible, use of summary charts showing their links).

## 2.7 EVALUATION RECOMMENDATIONS

**C12** Evaluation recommendations are generally coherent with the conclusions they derive from. However, the relationship between them should be reflected more clearly; and their specificity must be improved.

A linear and simplistic conceptualisation of processes usually results in problems. In the case of evaluation, the traditional concept considering it as the last stage in the process of public policy design took a too restrictive view, which did not understand its full potential to improve public policy actions before, during and after its application. In the NSRF context, we can see that the fact that recommendations are the last stage of the evaluation process could be affecting their development and practical focus. This is so, above all, because of the delays accumulated in launching evaluations and by the extension of the time periods initially established for field work in order to optimise the participation of key stakeholders. All this results in that the time available to carry out a comprehensive processing of relevant data and findings in order to draw accurate specific conclusions may be considered as too short.

In turn, the evaluators pointed out that the shortcomings which, in spite of the progress made, can be still detected in information systems lead to devote too many resources to collect and systematise information. This results in evaluation exercises being more descriptive than it would be desirable and leads to an imbalance with the proposal component that comes into being in the recommendations.

### RECOMMENDATIONS

**R15** The importance attached to establishing coherent, specific and practical recommendations in the Evaluation System of the NSRF context should be reinforced.

- **R15.1** Theoretical-methodological guidelines on the expected attributes of recommendations should be developed in the General Guidelines and the Terms of Reference, making emphasis if possible on their focus and practical implementation and on the fact that specificity shall be a requirement for the contracting institution to accept any Evaluation report.
- **R15.2** Evaluation Proposals should include a specific section for evaluation teams to state clearly the measures to be undertaken to ensure that recommendations are implemented and their commitments with regards to them.
- **R15.3** The Follow-up section should include different categories to classify the *reason for rejecting recommendations* ("unfounded", "application not described" or "excessive cost") so that it is possible to identify to what extent the reason for this is that they are not founded on the conclusions and on the data and findings obtained.
- **R15.4** All the Follow-up documents shall include *initiatives to be implemented* proposed by the Managing Authority, in order to ensure that accepted and partially-accepted recommendations are implemented, since it must be assumed that they are indeed a second level of implementation of recommendations (we could say their "administrative translation"). This is a task that the OP managers requesting an evaluation shall not delegate.

10

- **R15.5** One of the main tasks of evaluation managers and Steering Groups shall be to ensure that evaluations result in recommendations for improvement, coherent and specific in order to improve the programmes and policies being evaluated..

## EVALUATION STAGES: (III) DISSEMINATION AND FOLLOW-UP
**USEFULNESS:** Do evaluations meet the information and knowledge needs of stakeholders?

Analysis of the evaluation conclusions and recommendations as part the process to judge their accuracy has evidenced that evaluations answered to most of the questions posed in the Terms of Reference. To that end they met the minimum dissemination requirements established in the GEP. In average 73% of them were accepted by the Managing Authorities.

## 3.1 DISSEMINATION

**C13** The dissemination actions established in evaluations met the minimum GEP requirements; however it is clear that there is room for improvement.

Document analysis has allowed us to confirm that the requirements established to this regard have resulted in too generic guidelines in the General Guidelines on the dissemination focus evaluation processes should have as to qualify the political and public debate; moreover, the Terms of Reference are quite imprecise on this matter. These guidelines make emphasis on considering the different interests and needs of evaluation target groups although they do not establish how such response shall be guaranteed.

Both Evaluation Plans and Terms of Reference include the principles listed in the Guidelines and the GEP in a coherent manner, even though they carry on the same ambiguous nature and it does not seem to have specified sufficiently the requirements evaluations should have met to ensure them. Thus, we can state that evaluation reports met the specified requirements.

**C14** The role of the Executive Summary in the Evaluation System as a key element in its dissemination (especially in the General Guidelines and the GEP) does not match the attention given to it in Evaluation Plans, or the requirements specified in the Terms of Reference. This means that they are of no additional use either for the processes or for the Reports.

On the one hand, the lack of precision in the methodological and policy framework on requirements for evaluation dissemination and use is evident in the Executive Summaries; mainly when it can be stated that the group of Executive Summaries that do meet these requirements sufficiently are clearly subject to improvement and clearly insufficient to meet the dissemination function they should have. A further group of such summaries did not meet the requirements of the Terms of Reference. The requirements specified in the Terms of Reference for them fail to facilitate this purpose, both in terms of minimum content and probably regarding maximum extension too.

On the other hand, it should be taken into account that the Terms of Reference pose strict requirements on the Final Evaluation Report. They define precisely and accurately their methodological approach requiring a comprehensive account of all significant data and facts on the processes by means of a all techniques to make findings well-based and credible, and making specific recommendations that leave no doubt regarding their coherence and practical focus, Meeting all these requirements in the same document and accurately and at the same time making such documents easy to understand and appropriate for dissemination purposes is a very complex task, All the above has had an impact on the Executive Summaries which according to the Terms of Reference did not have such focus and therefore, evaluation teams did not devote more attention to them than merely including them as a separate paragraph to the Final report.

### RECOMMENDATIONS

**R16** Communication strategies planning processes should be reinforced as essential content of the Global Evaluation Plan and the specific Evaluation Plans, starting by a comprehensive involvement of stakeholders in their design and specifying the channels and formats to be used in order to ensure that they will meet the specific needs of each of them.

**R17** The purpose of the Executive Summary produced by evaluation teams should be more precisely defined, detailing their target groups, minimum contents and minimum length.

**R18** The implementation of supplementary dissemination and socialisation actions for target groups should be considered as independent from the actual evaluations themselves. They shall meet their different information needs, and provide customised products with the same structure and common image in order to ensure their quality.

**R.19** Basic evaluation documents should be submitted to NSRF Observatory.

**C14** The obligation to provide an answer to the recommendations established in the *Follow-up* procedures is a best practice of the System that is contributing to enhance promote the usefulness of evaluations, even though the malfunctioning detected in its application requires an optimisation of the process.

Considering the average response to recommendations (*Follow-up*) and their implementation status it has been found that in the prototype evaluation resulting from the seven that have been analysed 77% of them have been accepted (60% totally and 17% partially), 13% have been redirected and 14% have been rejected. There is a high variability in acceptation of the recommendations. In the POVT the acceptation ratio is 1 (100% were accepted) as compared to the NSRF (with 0.46 ratio). A similar pattern has been observed in the cases of INTERVIR, Mais Centro and O Novo Norte, with an acceptation ration around 0.7.

Taking into account the application ratio the recommendations, defined as the percentage of recommendations which have been actually implemented over the total number (implemented and underway), there were also great differences, ranging from 72% of implemented recommendations in INTERVIR to 0% of in the NSRF and the POVT.

**RECOMMENDATIONS**

**R25** Updating the *Follow-up* File and incorporating it into the Evaluation System in light of the findings of this Meta-evaluation.

**R25** Promoting Knowledge Management System to capitalise on the experience of and facts learned from evaluations, so that the dissemination of ideas across the Network can be enhanced in order for each of the Management Authorities to incorporate them into their operation.