Evaluation of Public Decentralised Cooperation Initiatives
Evaluation of Public Decentralised Cooperation Initiatives
Giorgio Mosangini

Bachelor in Political Sciences from the Université Libre de Bruxelles (ULB) with a Master’s Degree in Development Studies. He has 12 years’ experience in identifying, formulating, monitoring and evaluating international cooperation projects, including decentralised cooperation, in Latin America and the Mediterranean. Mosangini has produced different methodological guides, specifically on evaluating international cooperation. He has written articles and has run courses on evaluating projects. His work on public decentralised cooperation includes the production of a specific training module for the Catalan Agency for Development Cooperation and the Catalan Cooperation Fund.

Thanks to Beatriz Sanz for her advice in producing this guide.
Evaluation of Public Decentralised Cooperation Initiatives

Giorgio Mosangini
Contents

9 Preface

11 Introduction

13 1. Background and context of the guide
13 1.1. Local governments on the international scene
14 1.2. The new international cooperation agenda: aid effectiveness
16 1.3. The importance of evaluating decentralised cooperation

17 2. Evaluation concepts and approaches
17 2.1. Definition and purposes of evaluation
19 2.2. Evaluation types
22 2.3. Developments and trends in evaluation
24 2.4. Beyond types and approaches: a flexible perspective based on evaluation objectives and questions

29 3. Evaluation model
29 3.1. Classic evaluation criteria: the DAC model
39 3.2. Selecting own criteria: some proposals for evaluating DC
44 3.3. Beyond criteria: flexible models focused on evaluation objectives and questions

47 4. Management of evaluation

51 Methodology appendices
53 Appendix 1. Debate on quantitative and qualitative methodologies
55 Appendix 2. Description of some evaluation tools

59 Bibliography

64 Some websites upon which to find information on evaluation

Index of tables

20 Table 1: Main evaluation types
22 Table 2: Generations of evaluation
25 Table 3: Evaluation types prioritised for assessing DC
26 Table 4: Some limitations and difficulties faced by evaluation
30 Table 5: Some considerations regarding evaluation indicators
36 Table 6: Types of effects analysed by the criterion of impact
42 Table 7: The perspective of gender in evaluation
48 Table 8: Example of reference term contents
49 Table 9: Example of evaluation matrix
49 Table 10: Example of final evaluation report contents
55 Table 11: Examples of evaluation tools
The URB-AL III Programme Orientation and Coordination Office (OCO) is pleased to present a series of methodological guides resulting from the work and exchange with projects and lessons learned while the programme has been in progress. These guides cover a wide range of areas such as the monitoring, evaluation and communication of projects, the construction of local public policies in Latin America, the impact of these policies on social cohesion and the definition of city strategies with a focus on social cohesion.

The idea of producing a guide for the evaluation of public decentralised cooperation initiatives has arisen from OCO’s work of technical assistance in URB-AL III projects and from a need detected by the agents involved in local public management. It is a conceptual yet, above all, practical tool that initially sets in context the situation of local governments in the international system and in cooperation, and emphasises the importance of evaluation in fulfilling the principles of aid effectiveness in decentralised cooperation.

This guide also provides local governments with guidelines for showing results, elements to enhance transparency and accountability in management—on the grounds that evaluation can help provide responses and demonstrate achievements.

The tool presented here allows for evaluation of the contribution of a decentralised cooperation initiative or series of initiatives to local public social cohesion policies, and analysis of the extent to which intervention has generated, transformed and/or provided support to the subnational public policies that safeguard it.

Local governments are increasingly managing more initiatives in the international sphere and are present in areas previously only occupied by states and international organisations. Evaluation is therefore an essential element in developing new projects and strengthening, generating or innovating local public policies in different regions.

The guide thus also provides an overview of different approaches and trends in evaluation in order thereafter to give clear and precise bases for the five evaluation criteria of the Development Assistance Committee (DAC) of the Organisation for Economic Cooperation and Development (OECD), and three complementary criteria that complete the cycle.

We at the OCO hope that you enjoy reading and applying the guidelines provided in this guide practically, and that it should become your working tool for both present and future projects.

Jordi Castells,
Director of International Relations at Diputació de Barcelona and General Coordinator of the URB-AL III Programme Orientation and Coordination Office
This guide, Evaluation of Public Decentralised Cooperation Initiatives, is a URB-AL III Programme tool and an initiative of its Orientation and Coordination Office (OCO). URB-AL is a European Commission decentralised cooperation programme run among subnational European and Latin American authorities that responds to the political priority and interest of both regions to promote social cohesion.

The methodology is addressed to technical and policy teams of any Latin American or European local government that seeks to enhance local public social cohesion policies through its participation in decentralised cooperation initiatives. The guide therefore pursues a twofold objective:

/ To support subnational authorities in the accountability and evaluation of their decentralised cooperation initiatives, and of the impact thereof on local public social cohesion policies.

/ To encourage the collective learning of local governments and help identify and disseminate benchmark local public social cohesion policies.

The methodological contents are intended to provide the leaders of subnational authorities with a general awareness of approaches and models of evaluation and thus enable them to define proper evaluation objectives and questions to evaluate (generally on an external basis) their cooperation initiatives. The guide is also addressed to other agents involved in the evaluation cycle (external evaluators, target population of the cooperation initiatives and others).

The contents are divided into four chapters. The first deals with the background and context of the guide, and stresses the importance and relevance of having an evaluation methodology adapted to decentralised cooperation. Chapter two illustrates the main types and approaches and trends in evaluation, and places them within the context of cooperation led by local governments. Chapter three defines the guide’s evaluation model, establishes the criteria and other elements applicable in assessing decentralised cooperation initiatives and their impact on public policies, and stresses the core role of evaluation objectives and questions. Chapter four provides a summary of the main phases and instruments for managing the evaluation cycle. Lastly, the appendices deal with aspects of methodology and present some examples of evaluation tools.

We hope that the guide provides a useful tool for local governments to improve the quality of their cooperation initiatives and to encourage a culture of learning in their structures.
1. Background and context of the guide

This chapter features a brief description of two phenomena that make the availability of evaluation methodologies adapted to decentralised cooperation (DC) increasingly important: the growing presence of local governments (LGs)\(^1\) on the global stage and in cooperation, on one hand, and the international agenda of development focused on aid effectiveness and its impact on the cooperation system, on the other. It also looks at the importance of evaluating DC as a tool for improving quality and learning with a view to consolidating the contribution of cooperation initiatives to local public policies.

1.1. Local governments on the international scene

An initial item that is making DC evaluation methodologies increasingly necessary is the growing importance of LGs on the global stage in the last two decades. Although international relations have hitherto exclusively involved states and international organisations, LGs have gradually become regular agents in governance on a global scale. Very diverse dynamics have prompted the participation of sub-state governments in the international system. Globalisation of the capitalist system has accelerated in recent times and upon crossing borders has reshaped and incorporated local realities. In the light of growing global interdependence, LGs are acting on the international stage in order to influence the dynamics that affect them. Different transformations of an institutional kind have also driven the internationalisation of LGs. These include the erosion of the traditional role of the Nation-State and the transfer of power to supra- and sub-state levels.

In international cooperation, there has also been considerable growth in active LGs and the diversity of decentralised initiatives, ranging from aid and more traditional-type assistance to more horizontal and reciprocal cooperation initiatives based on the specific nature of the LG. In short, subnational authorities have emerged as agents in the international system and in cooperation. Acknowledgement by states and multilateral bodies of their role in development and governance has now therefore largely been firmly established.

---

\(^1\) The term local government (LG/LGs) will be used throughout the text to refer to very diverse sub-state authorities such as municipal, city, provincial or regional councils.
1.2. The new international cooperation agenda: aid effectiveness

The appearance of LGs in cooperation has been occurring against a background of profound change in the agenda and system of international cooperation for development. Since the mid-nineteen nineties, in a process led by the Development Assistance Committee (DAC) of the Organisation for Economic Cooperation and Development (OECD), governments and multilateral bodies that are active in the cooperation system have developed an agenda of international cooperation for development addressed to increasing the effectiveness of aid. This evolution occurred in a prevailing environment of in-depth debate on the real impact of cooperation.

The new agenda, which is focused on aid effectiveness, has been taking shape for over a decade, mainly at different international forums that have approved agreements and plans of action that specify strategies and instruments for raising the quality and effectiveness of the system.\(^2\) The different meetings and agreements notably include the 2005 Forum at which the Paris Declaration on Aid Effectiveness was approved. The Paris Declaration, which was signed by OECD member countries, multilateral development agencies, and over fifty aid recipient countries, is an international agreement that establishes the main bases for increasing the effectiveness and quality of aid for development. The text establishes five principles for smart aid (OECD 2008):

/ Ownership: partner countries establish their own policies and set their own development strategies.

/ Alignment: donor countries bring support in line with the development strategies of partner countries.

/ Harmonisation: donor actions are coordinated, transparent and effective.

/ Managing for results: results-oriented management of resources and decision-making.

/ Mutual accountability: donors and partner countries are accountable for development results.

Managing for Development Results (MfDR) is one of the key concepts associated with the search for greater effectiveness. In response to criticism of a lack of impact on the aid system, MfDR is a general management strategy geared to achieving better performance and specific and demonstrable results (for further information see, for example, UNDP 2009). Management systems are required to be able to demonstrate measurable and

---

meaningful results in the development conditions of target populations in order to justify the use of funds. MfDR also has an impact on evaluation, which it focuses on the intended results and impact of intervention, rather than on input, activities, processes or products.

The international cooperation agenda focused on effectiveness and its principles has shaped a new “international aid architecture”. This has prompted significant changes in priorities and in management mechanisms and instruments; for example, the appearance of country strategy-centred aid categories: Poverty Reduction Strategies (PRSs), Sector Policy Support Programmes (SPSPs), etc.

Neither has the effectiveness agenda been free of criticism. There follow two limitations arising from its implications in the matters dealt with in this guide. First, the new agenda is criticised for considering states to be the sole players of the cooperation system, and for paying insufficient attention to sub-state players (and to civil society). Application of the effectiveness agenda would therefore be contributing to strengthening the decision-making capacities of central governments while marginalizing the role of LGs and there would be a risk of weakening the decentralisation processes in progress (UCLG 2009). The efforts of subnational agents to “localise” and reappraise the effectiveness agenda have, however, also led to progress at summits and in international agreements. Since the Accra Forum there have therefore been advances and growing recognition of the specific role of LGs. Second, the effectiveness agenda has also been reproached for an excessively technical approach that does not adequately tackle the structural factors of development and of global inequalities. The ineffectiveness of cooperation is probably less a result of the weaknesses of the aid system (quantitatively marginal) than of other fundamental dynamics of the global system and their impact on the living conditions of populations that include flows of finance, inward investment, international trade, tariff policies, migratory flows, external debt, and military operations. If fundamental dynamics of the global system continue to hinder the fulfilment of targets to reduce poverty and inequalities, the effectiveness of the cooperation system is unlikely to improve.
1.3. The importance of evaluating decentralised cooperation

Against this new background, evaluation has become essential, as it is the main tool for answering questions on the effectiveness of development measures. The role of evaluation has therefore been stressed at different international forums that have shaped the aid effectiveness agenda. The final declaration of the Busan Forum lists transparency and accountability among the essential principles of the cooperation system. The economic crisis and budgetary cuts in cooperation also meanwhile exert pressure for greater accountability and for demonstrable results.

In this new context, DC cannot neglect the matter of evaluation. The effectiveness agenda requires LGs and DC to be able to prove their results, their added value, their specific contribution and comparative advantages. Evaluation can help DC to find answers to these requirements and demonstrate their contributions and specific benefits.

DC should not, however, be limited solely to demonstrating results, but rather has much more to offer and to contribute insofar as the aid effectiveness agenda is concerned. LGs therefore defend a specific approach in applying the principles of the effectiveness agenda that not only encourages ownership and alignment with regard to the plans of central governments, but is also based on LG strategies and on decentralisation, democratisation and citizen participation processes (see, for example, UCLG 2009 and Various Authors 2011). In this process, evaluation can also be a tool with which to make the strategies and specific contributions of DC visible.

Over and above renewed demands for accountability, DC and LGs may find in evaluation a key tool for learning and improving quality. As relatively recent agents and cooperation types in the system, they have even greater need to assess the quality of their strategies and intervention properly so that they may improve future initiatives.

Lastly, evaluation may also provide a source of legitimacy for the DC approach underlying the guide or, in other words, DC as a tool for strengthening local public social cohesion policies, based on horizontal and lasting relations among sub-state players. Having a clear approach helps the evaluation to enhance and make this visible and, therefore, fulfil the function of strengthening the legitimacy of LGs and DC in development policies.
2. Evaluation concepts and approaches

This chapter will feature a definition of evaluation and a review of the prominent intentions and purposes, and the main types of evaluations that currently exist. It will also deal with the main developments and trends that have shaped evaluation. We shall likewise indicate the perspective of the guide and detail what approaches and types are considered best suited for evaluating decentralised cooperation to improve local public policies. Emphasis shall meanwhile be placed on the fact that the core elements of any evaluation are its objectives and the questions that guide it. Types and approaches should therefore always be shaped as suitably as possible in order to respond to the ends pursued in the evaluation. Lastly, some difficulties and limitations that affect evaluation work will be mentioned.

2.1. Definition and purposes of evaluation

In cooperation, evaluation is understood to be the systematic and objective assessment of the design and execution of an initiative, and of its consequences on the lives of the people for whom it is intended. It is defined as a systematic and objective process insofar as it adheres to approaches and methods that are transparent, acknowledged and validated by the evaluating community and in accordance with international quality standards. Evaluation is also impartial and independent from interests associated with the management or political control of the initiative being evaluated. Lastly, evaluation must be of use for the agents involved in cooperation. All this provides credibility and legitimacy to the evaluation process as a whole and, therefore, to its findings and conclusions. In these circumstances, evaluation allows for answers to questions associated with the value or the importance of the development initiatives and their impact on populations.

The basic purpose of evaluation is to improve cooperation quality. In order to do so, evaluation should have two basic features: accountability and learning, which are briefly described below:

Accountability. Evaluation provides information on the general performance of initiatives and disseminates the results achieved, both in quantitative and in qualitative terms. Although very often demands for justification of the use of resources and attainment of results come from donors, the addressees of accountability can be very diverse (financing bodies, executing institutions, recipient bodies and partners, participant population, citizenry in general both in donor and in recipient countries, policy decision-makers, and other agents in cooperation, etc.). Evaluation can therefore help LGs to inform the public of the impact of cooperation relations and of whether they have helped to improve local public policies and how they have done so. Accountability prompts greater transparency in cooperation and in the local policies that have backed it and, therefore, increases its credibility and
legitimacy.

Learning. Commitment to learning is the purpose of the evaluation most closely associated with improving quality. Assessing what works and what does not allows for decisions geared to improvement. This partly prompts improvement in initiatives and cooperation policies by reorienting certain aspects thereof or by improving continuity or similar initiatives. It also encourages improvements in institutions and in the skills of the agents involved, from management mechanisms to elements of strategic planning. On the basis of evaluation processes, LGs can thus improve the quality of cooperation initiatives, their impact on local policies and at the same time strengthen their institutional powers for implementing them. Notwithstanding analyses focused on a specific initiative or actor, comparative evaluations that are focused, for example, on a specific theme or concrete sector provide for more easily generalised conclusions and summary recommendations, which may encourage learning among a large number of agents and thus prompt, for example, the dissemination of benchmark public social cohesion policies. Lastly, the purpose of learning from the evaluation is very closely associated with its utility, insofar as it can generate conclusions and recommendations that are efficiently incorporated in the institutional practices and realities of the agents to whom it is addressed.

In accordance with the two purposes of evaluation in cooperation briefly analysed above, this guide therefore pursues a twofold objective:

/ To support LGs in accountability and evaluation of their DC initiatives

The guide therefore provides methodology for evaluating public DC initiatives, and the contribution thereof to local public social cohesion policies.

/ To encourage collective learning

The methodology is also intended to encourage the exchange of experiences and public dissemination of DC best practices, as means for improving the quality and collective learning of agents in local cooperation.
2.2. Evaluation types

Different types of evaluation have been established in accordance with several classification criteria (purpose, moment in time, players involved, object, promoters, etc.). Table 1 shows a summary of the main resulting evaluation types. Table 3 of point 2.4 specifies the types that, to a greater or lesser extent, are suited to the needs of DC.
Participants involved

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Evaluation type</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Formative</td>
<td>Evaluation of an initiative throughout its execution process, focused on learning and on improving the design and performance of actions or of future initiatives.</td>
</tr>
<tr>
<td>External</td>
<td>Summative</td>
<td>Evaluation performed upon completion of the initiative and focused on accountability, on interpretation of the results and ultimate impact achieved and on decision-making with regard to continuity initiatives.</td>
</tr>
<tr>
<td>Mixed</td>
<td>Developmental</td>
<td>Ongoing evaluation (the evaluator is just another team member), focused on accompanying innovation processes in organisations and intervention in complex environments.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moment in time</th>
<th>Evaluation type</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex ante or prior</td>
<td>Evaluation of the design of an initiative prior to its implementation.</td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>Evaluation performed during the execution of an initiative, focused on improving its start-up.</td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>Evaluation performed upon completion of an initiative to assess its results.</td>
<td></td>
</tr>
<tr>
<td>Ex post</td>
<td>Evaluation performed some time after completion of an initiative to assess its impact.</td>
<td></td>
</tr>
<tr>
<td>On evaluability</td>
<td>Evaluation performed to determine the feasibility and relevance of an evaluation prior to its execution.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants involved</th>
<th>Evaluation type</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Evaluation is performed by people involved in executing the initiative. It has the advantage that the people who perform the evaluation are very familiar with the initiative and with the agents involved. It is, however, harder to demonstrate impartiality and independence, which may affect the credibility of the evaluation or even yield risks of conflicts of interests. Lastly, the people involved in executing the initiative do not necessarily have knowledge and experience in evaluation.</td>
<td></td>
</tr>
<tr>
<td>External</td>
<td>Evaluation is assigned to external agents who are not involved in executing the initiative. It features the advantage of impartiality and independence, and specialist knowledge and previous experience in evaluation. Both elements give greater credibility and legitimacy to the evaluation’s findings and conclusions.</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>Evaluation is the responsibility of a mixed team that includes representatives of the executor institution and external evaluators.</td>
<td></td>
</tr>
<tr>
<td>Participatory</td>
<td>The evaluating team comprises representatives of the communities in which the initiative is being executed. The evaluating personnel act as a facilitator of the process. The methodologies used ensure that the target population takes part in all phases of the evaluation process and that its results are useful to them. Local groups therefore agree on both the evaluation objectives and questions and on the use given to the conclusions and recommendations. By their very nature, participatory evaluations tend to focus more on analysis of processes and of changes in agents than on the results obtained.</td>
<td></td>
</tr>
</tbody>
</table>

Evaluation responds to the needs and the interests of the players involved in the initiative: donors and/or executors. Their guidelines determine the work of the evaluation teams when establishing objectives, methodologies and uses of the evaluation.

<table>
<thead>
<tr>
<th>Classification criterion</th>
<th>Evaluation type</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Formative</td>
<td>Evaluation of an initiative throughout its execution process, focused on learning and on improving the design and performance of actions or of future initiatives.</td>
</tr>
<tr>
<td>Summative</td>
<td>Evaluation performed upon completion of the initiative and focused on accountability, on interpretation of the results and ultimate impact achieved and on decision-making with regard to continuity initiatives.</td>
<td></td>
</tr>
<tr>
<td>Developmental</td>
<td>Ongoing evaluation (the evaluator is just another team member), focused on accompanying innovation processes in organisations and intervention in complex environments.</td>
<td></td>
</tr>
<tr>
<td>Ex ante or prior</td>
<td>Evaluation of the design of an initiative prior to its implementation.</td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>Evaluation performed during the execution of an initiative, focused on improving its start-up.</td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>Evaluation performed upon completion of an initiative to assess its results.</td>
<td></td>
</tr>
<tr>
<td>Ex post</td>
<td>Evaluation performed some time after completion of an initiative to assess its impact.</td>
<td></td>
</tr>
<tr>
<td>On evaluability</td>
<td>Evaluation performed to determine the feasibility and relevance of an evaluation prior to its execution.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Main evaluation types
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It is focused on evaluating the quality of the initiative's design, analysis of its relevance, and whether the logic of intervention suits the envisaged results and objectives, etc.</td>
<td>It is focused on evaluating the results envisaged for the initiative.</td>
<td>It is focused on evaluating a project, programme or specific cooperation policies, assessing its quality as a whole in accordance with diverse evaluation criteria.</td>
<td>It is focused on evaluating a specific instrument (for example, evaluation for different initiatives of technical assistance tools, support for budget items, thematic exchange networks, etc.).</td>
<td>It is focused on evaluating the impact of one or several initiatives or, in other words, on determining its/their effects and results as a whole in context, whether envisaged or not, or positive or negative.</td>
<td>It is focused on evaluating the process involved in executing the initiative, with emphasis on the impact on the agents and the institutional changes. To do so it analyses elements such as personal and institutional relations, decision-making procedures, improving skills, management mechanisms and agents' perceptions regarding the experiences.</td>
<td>Thematic evaluations are characterised by their comparative nature. They are based on evaluation of specific themes with regard to different initiatives. The content of comparative evaluations can be very diverse: geographical (countrywide, regional or worldwide), cross-disciplinary, institutional, strategy-based, intervention sector-based, etc. evaluation. In this guide, they can be joint valuations on the impact of a series of DC initiatives that analyse (among other things): LG financial and tax autonomy; citizen participation in the management of local public policies; support, from central governments, for decentralisation and LG autonomy; or the strengthening of a gender-oriented approach in local public policies. Comparative thematic evaluations help respond to the challenge raised by the multiplication of information generated by evaluation and favour learning among an extremely broad range of agents, as they provide conclusions and summary recommendations, applicable to different organisations and contexts.</td>
<td>This is focused on evaluating evaluations themselves and on the evaluation system. Meta-evaluations generally analyse a large number of evaluations, and assess their quality with regard to specific aspects (relevance, methodological rigour, utility and application, etc.). In addition to summarised dissemination of the knowledge accumulated by evaluations, meta-evaluations also generate transparency with regard to the evaluation system itself, demonstrate its limitations and strengthen its credibility and legitimacy.</td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>It refers to evaluations developed by the donor or the executing body.</td>
</tr>
</tbody>
</table>
2.3. Developments and trends in evaluation

As in any other area, evaluation is dynamic and has changed and developed over time. Texts on evaluation often feature reference to four basic generations of evaluation that indicate developments in prevailing approaches and perspectives. Although in practice the generations overlap in many aspects, the outline helps to understand the processes of change that have occurred.

The current result of the influence of successive generations is the coexistence of a host of evaluation approaches and perspectives. Changes in evaluation orientations also reflect the evolution over time of prevailing scientific

<table>
<thead>
<tr>
<th>Generation</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st generation</td>
<td>Evaluation is mainly focused on the quantitative measurement of products and results of initiatives to justify investment. It responds to the interests and perspectives of donors and whoever commissions it. The role of the evaluation team is mainly technical and focused on contributing quantitative methods and instruments of measurement.</td>
</tr>
<tr>
<td>(1900-1970)</td>
<td></td>
</tr>
<tr>
<td>Evaluation as</td>
<td></td>
</tr>
<tr>
<td>measurement</td>
<td></td>
</tr>
<tr>
<td>2nd generation</td>
<td>Evaluation is geared to measuring the effectiveness of initiatives and focused on analysis of expected results and objectives. Performance of cooperation initiatives is evaluated essentially on the basis of accountability and of transparency, and to improve initiatives. Evaluation still responds mainly to the interests and perspectives of donors. The role of the evaluation team is to describe what has happened in the project or programme with regard to the initial forecast.</td>
</tr>
<tr>
<td>(1980 onwards)</td>
<td></td>
</tr>
<tr>
<td>Evaluation as</td>
<td>Evaluation as description</td>
</tr>
<tr>
<td>description</td>
<td></td>
</tr>
<tr>
<td>3rd generation</td>
<td>Not only are the envisaged results and objectives evaluated, but the initiative itself (purposes, suitability, relevance, design, logic of intervention, etc.), unexpected or negative impact and processes are too. Judgement becomes a core element of evaluation. Qualitative tools are added to quantitative tools and an attempt to incorporate the perspective of target populations is made. The purpose of evaluation is essentially one of learning. Conclusions and recommendations arising from evaluations allow for improvement in processes and in intervention, and provide backup for decision-making. Evaluations still respond mainly to the interests of donors, although the impartiality of evaluation teams is now crucial. The evaluation team no longer provides solely methods of measurement and analysis, but an independent legitimate opinion about the initiative, its impact and its processes.</td>
</tr>
<tr>
<td>(1990 onwards)</td>
<td></td>
</tr>
<tr>
<td>Evaluation as</td>
<td>Evaluationasjudgement</td>
</tr>
<tr>
<td>judgement</td>
<td></td>
</tr>
<tr>
<td>4th generation</td>
<td>The approaches of evaluation are based on the participation of all agents, on the incorporation of their range of perspectives and on their learning. Evaluation is no longer limited to analysis of initiatives and their results and objectives, but is now extended, above all, to the study of processes (how and why objectives are achieved or not). Populations take decisions regarding the evaluation process and take part in each of its phases (from the definition of objectives to the application and implementation of recommendations). Methodologically, it requires greater flexibility and coordinates qualitative and quantitative methods. Evaluation no longer solely answers to the interests of donors, but must be oriented to suit the interests and perspectives of target populations. What is referred to as the 4th generation of approaches to evaluation includes a broad range of proposals, terminologies and trends (participatory, use-focused, empowering, democratic, etc. evaluation). Evaluation teams become facilitators with the essential role of ensuring that multiple perspectives and values are included, involving the greatest possible range of agents while mediating and negotiating their needs and interests.</td>
</tr>
<tr>
<td>(1990 onwards)</td>
<td></td>
</tr>
<tr>
<td>Participation in</td>
<td></td>
</tr>
<tr>
<td>evaluation</td>
<td></td>
</tr>
</tbody>
</table>

paradigms in different disciplines (see, for example, Izquierdo 2008). 

In recent years, several coexisting developments in evaluation can also be observed. One initial trend is a gradual change from evaluation of one-off initiatives to the performance of thematic and comparative evaluations. An increasing number of bilateral and multilateral agencies are focusing on thematic evaluations, by countries, of policies or strategies, while the number of agents who are focusing on evaluating specific projects or programmes is decreasing (Foresti 2007). Another parallel trend, prompted by the principles of alignment and harmonisation, is the growth in joint evaluations that facilitate the performance of comparative thematic evaluations and assessment of new cooperation instruments (support for budget lines, for sectoral policies, for country-specific strategies, etc.).

Evolution from evaluation focused on analysing specific projects and programmes to evaluation focused on processes, on institutional and systemic changes, very much suits the needs of the DC considered in this guide, the objective of which is to strengthen and generate local public policies. Hence, thematic and comparative evaluations respond more appropriately than analysis of specific initiatives to the methodological needs and analytical interests involved in the study of public policies. As mentioned above, evaluations that deal with cross-disciplinary matters can yield consolidated conclusions, recommendations and lessons learned that very much facilitate the collective learning of LGs. In accordance with the guide's purpose of helping to identify and disseminate benchmark models, comparative and joint evaluations can systematise best practices that encourage improvement in local public policies that contribute to social cohesion. The methodology described here therefore favours prioritising joint evaluation processes among different LGs focused on comparative thematic evaluations of DC that can encourage and improve collective learning of sub-state players with regard to local public policies.

Another trend observable in recent years is the emergence of articulation initiatives and efforts among institutional players and harmonisation of methodologies in

---

3 Hence the philosophical approaches of positivism and mechanicalism derived from physics and mathematics are seen to encourage an evaluative practice focused on exact reflection of reality, based on rigorous quantitative methods (experimental or quasi-experimental models). Realism (Westhorp 2011), on the other hand, questions whether it is possible to attain evidences and certainties, and leads to assessment, through qualitative and quantitative methods, of what works in which contexts and for whom. Lastly, constructivism sustains an evaluation focus nourished by disciplines such as anthropology or sociology that gives central importance to multiple perspectives and points of view on reality and reflects them mainly by means of qualitative research techniques.
the field of evaluation. For example, the growing number of areas for coordination and networking of the main agents in the international cooperation system: the OECD-sponsored DAC Network on Development Evaluation; the Evaluation Cooperation Group (ECG), established by multilateral development banks; the United Nations Evaluation Group (UNEG), which coordinates agents of the UN system; and international cooperation networks focused on the evaluation of impact such as the International Initiative for Impact Evaluation (3ie) and the Network of Networks on Impact Evaluation (NONIE). All these initiatives foster harmonisation of approaches and methodologies and encourage evaluation initiatives.

Lastly, it should be mentioned that the trends are not univocal and may sometimes seem contradictory. It is therefore necessary to relativise the emergence of an approach based on evaluations that are thematic, process-oriented and focused on more qualitative elements and the decline of ex post evaluations of specific initiatives based on quantitative methodologies. Recent years have indeed seen renewed interest in evaluations of impact based on experimental and quasi-experimental models. This recent trend has firm backing from international networks such as 3ie and NONIE, which are mentioned above, and may lead to the prioritisation of results and impact over processes, which is perhaps the result of the new aid effectiveness agenda. It should also be mentioned that although the importance of participation, of qualitative elements and of the incorporation of multiple perspectives in evaluation rhetoric and methodologies has been gathering strength for decades, participatory evaluations are, in practice, not widespread and are still in a minority.

2.4. Beyond types and approaches: a flexible perspective based on evaluation objectives and questions

Having reviewed evaluation types and approaches, a number of questions arise: which options are most suitable for evaluating DC and its contributions to local public policies? Which specific evaluation approach is adopted in the guide? We generally consider that evaluation of the impact of DC in local public policies should not be limited to accountability and requires more than evaluation approaches such as measurement and description (the 1st and 2nd generations briefly described above). Proper evaluation of developments in public policies therefore requires more than consideration of the envisaged results of cooperation initiatives, and should assess the effects as a whole (envisaged and not envisaged) of the initiatives on policies, and analysis of processes underway in the local reality and not only of specific projects and programmes. Evaluation must be focused mainly on learning and on improving quality, ensure the participation of different social groups and the inclusion of their perspectives, and make use both of
quantitative and qualitative methodologies (3rd and 4th generation). Analysis of different types of evaluation described in table 1 meanwhile shows that the characteristics of DC would

<table>
<thead>
<tr>
<th>Classification criterion</th>
<th>Evaluation type</th>
<th>Relevance in evaluating DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moment in time</td>
<td>Intermediate and final</td>
<td>Inadvisable because of its limitations in evaluating the effects of cooperation initiatives on public policies.</td>
</tr>
<tr>
<td></td>
<td>Ex post</td>
<td>Because the results and impact of DC initiatives on local public policies tend to be long-term, priority is placed on performing evaluations some time after completion of the initiatives for more suitable evaluation of their effects.</td>
</tr>
<tr>
<td>Participants involved</td>
<td>Internal</td>
<td>Inadvisable because of the complexity of the evaluation of impact on public policies.</td>
</tr>
<tr>
<td></td>
<td>External</td>
<td>The complexity involved in evaluating the impact of cooperation initiatives on local public policies (both regarding contents and from a methodological point of view) makes it advisable to perform external evaluations undertaken by evaluators with specialised knowledge and previous experience.</td>
</tr>
<tr>
<td></td>
<td>Participatory</td>
<td>Citizen participation legitimises local public policies and ensures that they are transparent and suited to the needs and interests of different collectives. Evaluating the contents of territorial social cohesion initiatives associated with DC and local public policies requires the participation of multiple agents and inclusion of the perspectives of varied social groups, and the use of qualitative methodologies. Encouragement of the highest possible level of participation is therefore advisable in all evaluations.</td>
</tr>
<tr>
<td>Object/content</td>
<td>Thematic</td>
<td>Without ruling out other types of content that may be necessary for the objectives considered when evaluating an initiative and its impact on local public policies, priority is given to the performance of comparative thematic evaluations. Hence, as has been repeatedly mentioned, their potential for strengthening the collective learning of LGs through the generation of conclusions and summary recommendations, which facilitate the identification and dissemination of benchmark local public social cohesion policies.</td>
</tr>
<tr>
<td>Promoters</td>
<td>Joint</td>
<td>LGs should try to perform joint evaluations. That would encourage combined efforts and resources to deal with complex evaluation questions, perform comparative thematic evaluations and generate summary conclusions that benefit the greatest number of LG and prompt collective learning. For the dissemination of best practices and lessons learned in accordance with the DC approach dealt with here, setting up a common evaluation fund among LGs could therefore be very useful.</td>
</tr>
</tbody>
</table>
favour the prioritisation of certain types of evaluations over others. Table 3 shows the main types that are most suitable for evaluating DC initiatives and their impact on local public social cohesion policies.

Notwithstanding these general orientations regarding the approaches and types more likely to be appropriate and suitable, it should be emphasised that the essential elements of any evaluation are the objectives and questions by which it is guided. What purposes are being pursued in the evaluation? What matters should it elucidate on? Why? What agents will use the knowledge produced? How will it be

| The viability and evaluability of evaluations | Since it is not possible to evaluate everything (because of limited time, resources, capacities to assimilate results, etc.), evaluations should be chosen according to relevance, feasibility and utility. |
| Methodological limits | In light of the diversity of DC experiences (multiplicity of agents, scales of intervention, themes, sectors, objectives, etc.), it is difficult to establish general validity methodologies. The contents of the guide should therefore be used as general guidelines, approaches and methods to be adapted to the objectives and questions of each specific evaluation. |
| The problem of attribution | One of the main difficulties faced in evaluation, and which will be dealt with in chapters 3 and 5, is determining whether or not a result or impact has actually arisen from the initiative. |
| Specific difficulties in evaluating DC and its impact on public policies | In addition to the problem of attribution, there are also other specific difficulties in evaluating DC and its influence on local public policies. First, evaluation in international cooperation has mainly been focused on evaluating projects or programmes. Study focused on the impact on public policies is less common and prompts greater methodological challenges. Evaluating the effect of DC initiatives on local public policies also involves a medium- to long-term timescale as in many cases the influences of intervention are likely to appear years after they are complete. For example, in cooperation initiatives that have accompanied decentralisation processes, it has been observed that against backgrounds of financial or political instability, effective initiatives have remained so after over a decade of work (OECD 2004). These timescales give rise to significant attribution problems in the light of causes that are diluted over time. The use of impact evaluations, the most complex and demanding insofar as human and economic resources and time are concerned, is advisable. One last difficulty as regards DC is associated with the lack of consensus as to the purposes and specificities of this cooperation type. As far as this guide is concerned, this difficulty can be overcome by the approach adopted in the URB-AL III Programme, which shapes their objectives: DC is intended to improve and generate local public policies that contribute to social cohesion and are based on horizontal cooperation mechanisms and on the mutual interest of LGs. Having a clear perspective about DC and its objectives very significantly facilitates the use of evaluation and enhances its utility. |
| The scope of evaluation and consistency of policies | Analysis of the results and impact of cooperation initiatives shows that often evaluations do not heed structural factors that condition the effectiveness and viability of cooperation, such as other international policies and relations. Studies focus on initiatives but scarce attention has been paid to evaluating global cooperation policies, their relation with other policies, and the consistency thereof with targets to reduce poverty and inequality (Unceta 2011, Izquierdo 2008). |
applied? What for? The evaluation types and approaches will be coordinated in accordance with these concerns with regard to learning and improving the quality of LGs.

Hence, although the guide points to more suitable specific approaches or types of evaluation, it mainly defends the central importance of the objectives and of the questions of evaluation. The purpose established by LGs in evaluating conditions all the other components of evaluation: approach, type, methodology, organisational mechanisms, application strategies and widespread implementation of recommendations, etc. Such a view of evaluation also involves consolidating a culture of evaluation. This means that LGs must integrate evaluation and learning in their mechanisms and structures and use the knowledge yielded by evaluations to improve quality.
3. Evaluation model

This chapter features a description of different evaluation criteria (the five proposed by the DAC model and three additional criteria), and their application to DC. It will also include fresh emphasis on the central importance of the evaluation objectives and questions that should guide the selection of criteria and definition of other elements to characterise the evaluation model chosen. When dealing with criteria it is therefore always important to remember that they are general guidelines for reflection that should be determined in order to provide an appropriate response to the specific questions of each evaluation.

3.1. Classic evaluation criteria: the DAC model

In the previous chapter it was mentioned that the basic elements of an evaluation are the questions that guide it. The evaluation was also defined as an instrument for answering these questions by systematically assessing an initiative and its consequences on the life of the people with whom it is being implemented. Bearing these elements in mind helps to understand what the evaluation criteria consist of.

The criteria are concepts that have been defined with a view to organising and answering the questions of any evaluation type. They could be described as representing the dimensions of quality commonly accepted in international cooperation, or, in other words, as points of view that provide a basis for assessing the quality of the cooperation initiatives. The criteria are also very closely associated with the systematic characteristics of the evaluation. Unlike a judgement or an opinion, the previous chapter showed that evaluation involves systematic assessment or, in other words, is based on internationally validated approaches and methodologies. The evaluation criteria are some of the elements that allow for systematic and transparent assessment, and therefore give credibility and legitimacy to the findings, conclusions and recommendations.

The criteria adopted and used by most public bodies and private institutions in the cooperation system are those put forward by the DAC of the OECD. They are five criteria that articulate international consensus and standards on quality in evaluation: relevance, efficiency, effectiveness, impact and sustainability. Although international development bodies and bilateral cooperation agencies do not adopt a uniform evaluation methodology, they do include the five DAC criteria as a core feature of their evaluation models. Notwithstanding this minimum consensus, specifying methodologies depends on each agent and on the proposals put forward by the evaluation

---

4 Some of the contents in this section have been adapted from Gascón and Mosangini 2009.
teams (generally external) in response to the objectives and questions of the evaluations.

The content of each of the five criteria are detailed below. Each section will suggest some examples of indicators for assessing DC initiatives and their impact on local public policies. The indicators are used to make the criteria operational. While the criteria allow for systematisation and organisation of the evaluation questions, the indicators specify for each criterion the contents of the responses that should be sought and point to the necessary sources of information. The sample indicators here are of a general nature and intended to help understanding of the contents. They do not therefore meet the quality requisites required of the indicators (see table 5) and are not directly applicable. Specific indicators must be defined for each specific evaluation.

### Relevance

**Definition:**

Relevance evaluates the extent to which the initiative is suited to the needs, skills and priorities of local agents, in accordance with the policies of donors.

**Key questions:**

/ Is the initiative suited to local needs, skills and priorities and those of donors?

/ If there are changes in needs and priorities, how has the initiative been adapted?

Relevance analyses the quality of initiatives according to the extent to which they are suited to the context in which they operate. This involves contrast of the cooperation initiative with its environment: diversity of agents, perspectives, needs and interests, policies, institutions, etc. These elements influence the initiative’s legitimacy, the target populations, the problems tackled and the chances of achieving the envisaged results and objectives.

---

**Table 5: Some considerations regarding evaluation indicators**

<table>
<thead>
<tr>
<th>A distinction is usually made between quantitative (expressed quantitatively and numerically) and qualitative indicators (expressed by evaluations and descriptions), and other direct indicators (which express the variable to be measured) and indirect indicators (which express the behaviour of a variable by measuring another that is less complex or less difficult to measure).</th>
</tr>
</thead>
<tbody>
<tr>
<td>To have quality status, an indicator must fulfil a series of characteristics such as being specific, measurable, attainable, relevant, and having time or economic parameters.</td>
</tr>
<tr>
<td>Indicators are sometimes categorised according to their purpose in the evaluation (in accordance with the object evaluated: indicators of process, result, impact, etc.; in accordance with the theme evaluated: social, gender, sectoral, geographical, etc. indicators).</td>
</tr>
</tbody>
</table>
The relevance of an initiative measures whether it is suited to: the needs, priorities, outlooks and lifestyles of the populations in the regions in which it is being implemented; the policies and strategies of local, regional and state institutions; and the policies and strategies of donors. If it is neither useful nor coherent for these agents, a cooperation initiative should not be run. Relevance also analyses whether the proposed technical solutions are suited to the problems identified and to the skills of local agents and/or executors. Because of its very nature, the criterion focuses its attention on a priority basis on the aspects and phases of planning, identification and design of initiatives.

Because context changes constantly, analysis of relevance should however be dynamic and assess initiatives’ adaptation capacities.

It is important to note that relevance has different dimensions that may be contradictory without necessarily affecting the overall relevance of the initiative. For example, the relevance of an initiative does not always mean it is suited to state policies, but that it has taken them into consideration for the design of its intervention strategies. Imagine an LG that opts to strengthen its autonomy and powers through participation in a DC initiative despite central government reticence with regard

<table>
<thead>
<tr>
<th>Examples of general indicators of relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>To evaluate DC activities</strong></td>
</tr>
<tr>
<td>The initiative objectives have been defined in order to incorporate the needs and interests of different social groups.</td>
</tr>
<tr>
<td>Mechanisms of execution have been designed in accordance with LG institutional realities and capacities.</td>
</tr>
<tr>
<td>The initiative and local policies it supports are relevant to state policies or, conversely, the design of the initiative has incorporated strategies to deal with possible contrasting positions.</td>
</tr>
<tr>
<td>The cooperation initiative responds to the shared and reciprocal interests of the Latin American and European participant authorities and is based on medium- and long-term collaborative relations.</td>
</tr>
<tr>
<td>The initiative has been aligned with the main local, national and international instruments to foster gender equality (CEDAW, etc.).</td>
</tr>
<tr>
<td><strong>To evaluate the contribution to local public policies</strong></td>
</tr>
<tr>
<td>The public policies prompted by the initiative are suited to LG institutional powers.</td>
</tr>
<tr>
<td>The public policies that the cooperation initiative has helped generate respond to the LG agenda.</td>
</tr>
<tr>
<td>The local public policies it supports incorporate needs and interests of social groups that are discriminated (for reasons of sex, socio-economic status, membership of ethnic minorities, disability, etc.).</td>
</tr>
<tr>
<td>The initiative has prompted the LG to provide public services suited to the needs and requirements of the public.</td>
</tr>
</tbody>
</table>
to the decentralisation processes. In this event, the relevance of the initiative does not stem from its suitability to state policies, but from evaluating whether the design of the initiative has taken these tensions into account and whether suitable strategies have been established to deal with the situation.

Efficiency

Definition:

Efficiency measures the relation between invested resources (economic, human, technical, material and temporal) and the results obtained.

Key questions:

/ Could the same results have been attained using fewer resources or in less time?

/ Could the resources used have attained more and/or better results?

Efficiency is an essentially economic analysis that assesses whether the resources invested by the initiative have been used optimally and have prioritised less costly options.

It analyses processes through which different types of resources mobilised by the initiative (financial, human, material, technical, time, etc.) have been turned into results. It studies whether the same results (both quantitative and qualitative) could have been attained with fewer resources, with greater optimisation or in less time. It also analyses whether alternative use of different resources assigned to execution of the initiative could have yielded more significant quantitative or qualitative results. The maximum degree of efficiency of an initiative would therefore be attained when the value of its results exceeds any alternative use of the resources. The criterion, by nature, focuses on certain elements of intervention such as activities, the budget, the performance schedule, management procedures and results.

Efficiency is never absolute and is always established in comparison with alternative assignations of resources or other ways of attaining results. The criterion is usually measured with tools from economic science (cost-benefit analysis for example), although it also requires analysis of a more qualitative (assessing whether the initiative has resorted to local personnel and locally acquired resources, for example) or political (an initiative may involve the purchase of local, albeit more expensive, goods to favour local employment) type.

Lastly the criterion has some limitations, which must be considered when applying it. On one hand, efficiency depends on effectiveness. What is essential is that an initiative should be efficient or, in other words, it should achieve the objectives and results proposed, and thereafter establish whether it has done so in the most efficient way possible. Use of tools such as cost-benefit analysis may also introduce biases
with regard to future generations and environmental sustainability, as they tend to incorporate neither the interests of future populations adequately nor the real costs in ecological terms. Lastly, positive analysis of efficiency, such as an increase in a service, can hide inequalities in access (according to gender or social differentiation criteria, etc.).

### Effectiveness

**Definition:**

Effectiveness measures the extent to which the initiative has attained its objectives and results.

**Key questions:**

- Have the established objectives and results been attained? Give reasons.
- Could the resources used have attained more and/or better results?

Effectiveness measures the extent to which the initiative has fulfilled or will fulfil its objectives and results. More specifically and in accordance with the vertical logic of intervention, it determines whether the initiative has fulfilled the results expected and consequently achieved its specific objective. It also evaluates the main reasons regarding how and why the
results and objectives have or have not been achieved.

To do so the evaluation must measure results and objectives by applying quantitative and qualitative tools, using the instruments produced by the initiative itself (diagnoses, bottom line, indicators, verification sources, etc.). Upon verification of the extent to which results and objectives have been met, evaluation of effectiveness must attribute the effects to internal causes of the initiative and/or to external causes in order to judge the quality of intervention in fulfilling the established targets.

Analysis of effectiveness is very much conditioned by the quality of the design of initiatives. Limitations in the design of an initiative can seriously affect the ability to measure its effectiveness. There follow some examples of weaknesses in the design of an initiative that hinder evaluation of effectiveness: the initiative does not have clear results and/or objectives; it has neither quality indicators nor verification sources; it did not establish a bottom line or diagnoses to reflect the initial situation; the logic of intervention is incorrect, as the activities envisaged do not allow for proper attainment of the results, and these results do not guarantee fulfilment of the specific objective. Faced with such difficulties, the evaluation team is forced to generate information subsequently and even reformulate operational results and objectives for the evaluation process.

Lastly, when measuring effectiveness it is important to keep the process focused on the target population because results and objectives must be addressed to improving its living conditions. To do so, it is crucial to bear in mind the differences that exist within the population under consideration (information should be broken down by sex, socio-economic status, age or ethnic group) in order to detect any bias in access to the results of the initiative. The involvement of different groups, from the
outset of the initiative and in every phase of the intervention cycle, is an element that conditions effectiveness.

Impact

Definition:

Impact is the series of effects (whether positive or negative, expected or not) produced by the initiative on the population and the environment.

Key questions:

/ What effects has the initiative had on the population and the context? Have they been positive or negative and were they expected or not?

/ Can the effects be attributed to other causes?

/ What would have happened if the initiative had not been run?

As indicated in its definition, impact covers all the effects of an initiative on the population and the general context. Effects are understood to be all the changes produced by an initiative. The criterion is very broad and involves consideration of all the initiative's possible consequences. Table 6 shows the main types of effects an initiative may have.

Analysis of impact initially requires identification of all the changes that have occurred in the reality under consideration, upon comparison with the situation prior to intervention. The information provided by the initiative itself may help in the task (bottom line, initial diagnoses, indicators, etc.) although, as stated previously, this will not always be available. Although the tools of initiatives meanwhile tend to be addressed to measuring the expected results, they are not much help in appreciating the unexpected or undesired effects. Upon determining the main changes in the population and the environment, the evaluation should then help to establish whether or not there are causal links between these and the initiative. By excluding the effects produced by external causes and systematising the changes attributable to the initiative their impact is demonstrated. It can be very hard to establish categorically that the initiative has yielded specific changes and effects, as phenomena tend to arise because of a wide range of causes and factors. Separating the effects of the initiative from external effects is not always possible. As explained in methodology appendix 1, many authors argue that to overcome the difficulties of attribution and for evaluations to be truly considered evaluations of impact, they need to be performed under experimental conditions of randomisation. Experimental or quasi-experimental models allow for comparison of two population groups: one of which has been subject to the initiative while the other has not.
Having a control group with characteristics similar to the target population that has not taken part in the initiative (counterfactual) should provide sufficient grounds to allow the effects attributable to the initiative to be separated from those with other causes. With such methodological requirements, however, few evaluations in cooperation can be considered to have an impact. In most situations, evaluations opt pragmatically to combine quantitative non-experimental methods and qualitative methods, in search of a multidimensional understanding of the impact and of the processes.

In addition to the uncertainty arising from attribution there are the demands associated with time, knowledge, methodology and economic and material resources with which evaluations of impact are performed. This situation makes the criterion extremely complex to assess. Such difficulty makes evaluations on an initiative’s general impact that tackle its effects as a whole inadvisable as they are too ambitious. It is usually more useful and viable to perform evaluations of impact based on a specific theme in order to prompt learning or improvement. Hence, evaluations could be performed on the impact of a series of DC initiatives.

<table>
<thead>
<tr>
<th>Type of effect</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected and unexpected</td>
<td>Impact not only assesses the expected results and objectives, but also effects that were not envisaged in the intervention logic. It therefore analyses the unintentional consequences and implications yielded by the initiative.</td>
</tr>
<tr>
<td>Positive and negative impact</td>
<td>Impact is not restricted to assessing the positive effects; it also analyses the unwanted negative consequences or the unwanted effects an initiative may have had.</td>
</tr>
<tr>
<td>On the population</td>
<td>Impact not only studies the effects of the initiative on the target population, but also analyses its consequences on the population as a whole and evaluates any bias and possible differentiation in impacts associated with sex, socio-economic status, age, membership of ethnic minorities, or any other specific characteristic of a group under consideration.</td>
</tr>
<tr>
<td>On the environment</td>
<td>Impact analyses the effects on the environment and context, defined in a broad sense: political, social, economic, cultural, environmental, and institutional, etc.</td>
</tr>
<tr>
<td>In the short-, medium- and long-term</td>
<td>The impact of an initiative is generally discerned in the medium- to long-term (the terms ‘evaluation of impact’ and ‘ex post’ are therefore sometimes synonymous); evaluation of impact analyses effects in all possible timescales.</td>
</tr>
<tr>
<td>Direct and indirect</td>
<td>Impact is not limited to the initiative’s most apparent and clearest effects, but also features analysis of possible indirect or secondary effects.</td>
</tr>
<tr>
<td>Locally, nationally or regionally</td>
<td>Impact analyses the effects of the initiative that may have arisen on any geographical or regional scale, locally, regionally, nationally or internationally, on a micro (on people’s living conditions) or on a macro level (in sectoral policies).</td>
</tr>
</tbody>
</table>

Table 6: Types of effects analysed by the criterion of impact
on decentralisation processes, on the establishment of citizen participation mechanisms, on the generation of tools to mainstream public policies among different departments or subject areas, on women’s participation in politics, or on the reduction of regional inequalities.

Sustainability

Definition:

Sustainability assesses whether the benefits of an initiative are likely to continue once it has been completed.

Key questions:

/ Which of the initiative’s benefits are likely to continue in the long-term and which are not? Why?

/ Do the capacities and the context of the target population and of local institutions allow them to assume the results of the initiative as their own?

Sustainability analyses the chances that the benefits of a cooperation initiative will continue after its completion. It is basically associated with the capacities of target populations and of local institutions to assume continuity of the desired impacts (activities, results, services, infrastructures, cultural changes, intangible benefits, etc.). Since responsibility for continuity of the effects lies with local agents, analysis thereof must necessarily be focused on the population and institutions associated with the initiative and concentrate on aspects such as: the suitability of the intervention strategy to local realities; the participation of local agents in every phase of the initiative; the mechanisms of citizen participation envisaged to sustain the results; the technical and human skills of local agents, etc. The design of the initiative also influences sustainability. It is therefore necessary to analyse the extent to which viable and effective transfer mechanisms were envisaged when

<table>
<thead>
<tr>
<th>To evaluate DC activities</th>
<th>The initiative has yielded positive effects on LG capacities, powers and structures.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The initiative has yielded positive effects on the citizens’ living conditions and access to public services.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>To evaluate the contribution to local public policies</th>
<th>The positive impact of the initiative on local public policies has been possible because of the specificity and added value of the DC approach applied.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The effects of the initiative on local public policies have prompted positive impact for citizens and different social groups.</td>
<td></td>
</tr>
</tbody>
</table>
the initiative was designed in collaboration with local agents.

As sustainability largely depends on the context, evaluation of cooperation has involved attempts to systemise the most significant elements of an initiative’s environment that influence its viability. Hence definition of the following development factors: support policies (do local, regional and national policies provide continuity of impact?); institutional aspects (do local agents have the capacity to maintain the benefits and are these suited to their priorities?); socio-cultural factors (are the initiative and its effects in keeping with the characteristics of local socio-cultural systems?); economic factors (is the income necessary to maintain benefits guaranteed?); technological factors (do the technologies applied suit local conditions and local capacities?); ecological factors (do the initiative and the continuity of its benefits respect environmental sustainability?); gender factors (to ensure the initiative’s viability, have inequalities between sexes and mechanisms for balancing power and the distribution of benefits among men and women been taken into account?).

Sustainability is very often the Achilles’ heel of cooperation initiatives. DC is no exception and evaluations also tend to reflect considerable difficulties in this regard (OECD 2004). From the perspective of the guide, which contemplates DC focused on supporting local public policies, sustainability is even more strategic. In the guide, sustainability in DC consists in taking a step away from the elements involved in an initiative (activities, results, products, services, approaches, benefits, changes in values or procedures, etc.) and towards the consolidation of local public policies. If DC initiatives, which by definition are intended to improve LG policies, do not achieve sustainability, then there is no longer any point to them.

The viability of the policies implemented should therefore be a central feature of DC.

<table>
<thead>
<tr>
<th>To evaluate DC activities</th>
<th>From its design, the initiative has envisaged strategies to institutionalise the local policies encouraged.</th>
</tr>
</thead>
<tbody>
<tr>
<td>To evaluate the contribution to local public policies</td>
<td>Local public policies respond to the agenda and commitments assumed by LGs to citizens.</td>
</tr>
<tr>
<td></td>
<td>Measures have been implemented to encourage support by state policies for enhanced local policies.</td>
</tr>
<tr>
<td></td>
<td>The local public policies backed contemplate citizen participation mechanisms for review, monitoring and evaluation.</td>
</tr>
<tr>
<td></td>
<td>The capacities, powers and economic independence of LGs have been strengthened for it to sustain the backed local policies.</td>
</tr>
</tbody>
</table>
3.2. Selecting own criteria: some proposals for evaluating DC

The five criteria set out in the evaluation model proposed by the DAC and adopted by the cooperation system allow for systemisation of many of the questions and needs for information required to ensure quality evaluation of an initiative. The DAC criteria should, however, be considered neither exhaustive nor exclusive. When considering an evaluation, the use of supplementary criteria may be necessary to fulfil perspectives and needs for information. There are some elements to which the classic criteria do not pay sufficient heed and therefore may ultimately not be evaluated and analysed. The use of some criteria, moreover, does not rule out the use of others.

This section therefore features suggestions for some complementary criteria that we consider useful for in-depth evaluation of specific aspects of DC initiatives and their impact on local public policies. There follow some very brief descriptions of three criteria that in our opinion answer the needs for strategic information in order to fulfil the guide’s objectives. These criteria are coherence, coverage and innovation.

Coherence

Definition:

Coherence assesses the extent to which background policies encourage the achievement and sustainability of an initiative’s benefits.

Key questions:

/ Do the most significant policies with an impact on the context complement or run counter to the initiative’s benefits? Why?

Coherence is understood to mean the evaluation of the relation between policies of the context and the initiative’s results and objectives. This relation may affect the achievement and continuity of the expected results in different ways, ranging from reinforcing or upholding them over time, to hindering them or reverting them. This guide includes consideration of any policy (on a local, regional, state or multilateral scale) and not only development policies (other cooperation initiatives for example), but also any other existing policies: economic, trade, industrial, financial, environmental, diplomatic, military, migratory, etc. Policy here should, moreover, not be understood to refer exclusively to the public domain but could also include, for example, business strategies that affect the achievement or sustainability of the initiative’s benefits in the region.
The criterion leads to analysis that includes the degrees of interdisciplinarity attained in the initiative’s approaches and benefits. This involves assessing synergies, areas of coordination and articulations generated among different LG institutional areas with other sub-national governments or with regional, state or multilateral agents.

Lastly, coherence also involves evaluation of the suitability of the initiative and of background policies to concepts that enhance local independence, such as decentralisation (political, administrative, fiscal), subsidiarity (public policies are assumed by the most immediate authorities) and citizen participation (the population takes active part in defining and managing local development strategies).

Coverage

Definition:

Coverage assesses whether an initiative has reached all the target groups and the reasons behind any possible breaches of access.

Key questions:

Are there any differences between the initially envisaged target groups and those who have had access to the initiative’s benefits? Give reasons.

Coverage assesses whether specific groups of citizens targeted by the initiative have been subject to biases or barred access. Breaches in access to the initiative’s benefits may affect groups because of elements of

<table>
<thead>
<tr>
<th>To evaluate DC activities</th>
<th>The initiative is included in the LG strategic plans (municipal development plan or similar).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Execution of the initiative is backed up with the coordinated action of different LG departments and thematic areas.</td>
</tr>
<tr>
<td></td>
<td>The initiative has developed strategies to deal with contradictions among backed local public policies and other policies that affect the target population.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>To evaluate the contribution to local public policies</th>
<th>The initiative has helped other agents to modify policies or actions that generate imbalances and obstacles (political, economic, ecological, etc.) to starting up LG public policies.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The initiative has strengthened policies that encourage a framework to distribute institutional powers based on subsidiarity.</td>
</tr>
<tr>
<td></td>
<td>The initiative has helped to formulate policy proposals that stretch beyond the territorial scope of the LG and support acknowledgement of the role and specific characteristics of the sub-state authorities.</td>
</tr>
</tbody>
</table>
differentiation and exclusion that range from sex, socio-economic status, age, membership of ethnic groups, geographical location, skills, religion, ideology and job to networks of social relations. Groups included initially may be excluded while groups not envisaged within the target population can benefit from the initiative. Biases in access may affect the benefits of the DC initiative and/or local public policies and services supported thereby. Inequality in access can be explained by the fact that existing differentiations have not been taken into account in the design of the initiative, the population having been considered as a homogeneous whole, or because the cooperation initiative and/or the agents involved reproduce patterns of exclusion.

The coverage criterion studies elements such as the coverage rate (percentage of the target population that has had access to the initiative's benefits), coverage biases (common characteristics of groups included and excluded by the initiative) or accessibility (elements that are internal or external to the initiative that explain unequal access). Analysis can be systemised by establishing lists of local groups and of the initiative's benefits and results, contrasting the access of each group for each impact category and revealing possible breaches of access.

The coverage criterion is considered to be crucial in evaluation methodology focused on analysing local public policies, because public access and the elimination of inequalities depend on it.

### Examples of general indicators of coverage

<table>
<thead>
<tr>
<th>To evaluate DC activities</th>
<th>The initiative's benefits have reached all the envisaged target groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The initiative featured mechanisms for the participation of different social groups in its different phases (design, execution, monitoring, evaluation, etc.).</td>
</tr>
<tr>
<td></td>
<td>The DC initiative has reached the most underprivileged social groups; for example, initiatives focused on creating employment have had an impact on the groups most affected by unemployment, such as women or young people; projects that improved basic social services have prioritised population segments that were not benefiting from coverage by these services; etc.).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>To evaluate the contribution to local public policies</th>
<th>The initiative has improved the access of excluded and underprivileged groups (for reasons of sex, socio-economic status, age, membership of ethnic groups, etc.) to local public policies and services.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The local public policies backed help to reduce regional inequalities.</td>
</tr>
<tr>
<td></td>
<td>Mechanisms of citizen participation that reflect the needs and interests of different groups that compose the local citizenry have been created and consolidated.</td>
</tr>
</tbody>
</table>
of breaches of access are core elements of public management.

Innovation

Definition:

Innovation assesses whether an initiative features experimental elements and how these affect assessment of its quality and its potential replication.

Key questions:

/ Does the initiative feature experimental and innovative elements? How have they influenced its performance and quality?

/ Has the initiative generated new tools or strategies? Do they have the potential to be replicated?

Examples of general indicators of innovation

<table>
<thead>
<tr>
<th>To evaluate DC activities</th>
<th>The initiative has yielded lessons learned and innovative benchmark models for replication and application by LGs in other regions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>To evaluate the contribution to local public policies</td>
<td>The initiative has helped in the design and application of new local public policies, and of the instruments, services and programmes associated therewith.</td>
</tr>
<tr>
<td></td>
<td>The initiative has introduced innovations in local public policies and these have tended to improve their impact on social cohesion</td>
</tr>
</tbody>
</table>

Table 7: The perspective of gender in evaluation

The incorporation of gender as an evaluation criterion is relevant in any evaluation type. Indeed, the evaluation model should deal with gender both on a cross-disciplinary basis (in all criteria) and specifically (evaluation of the whole initiative from a gender perspective). Because space in this guide is limited, it will only feature some general questions that may facilitate the incorporation of a gender perspective in evaluation.

/ Have the diagnosis and design of the initiative included a gender perspective and the participation of women?
/ Does the budget contemplate resources allocated to reducing inequalities among men and women?
/ Has the initiative generated changes in the position of women with regard to men (economically, politically, culturally, etc.)?
/ Has the initiative increased women’s decision-making capacities?
/ Has the initiative attended the needs and interests of women?
/ Have the objectives and the results taken into account the inequalities of resources and power among men and women?
/ How do the products and results of the initiative benefit women and men?
/ Has the context made the participation of women difficult?
/ What impact has the initiative had on the relations of power and inequalities among women and men? Has it had any unforeseen or undesired impacts on women?

Compiled by author based on: Murguialday et al. 2008 and UNEG 2011.
As a general trend, evaluation criteria generally tend to penalise innovative initiatives and reward more traditional initiatives. Innovation, however, is an important element for change and, therefore, for generating processes of quality improvement. This is also true for DC. Initiatives that seek to improve or transform local public social cohesion policies very often need to incorporate innovative elements. The idea that innovation must be a characteristic of DC can even be defended. It is therefore an element that allows for exploration of new strategies and tools and the subsequent generalisation of successful experimental public policies in order to improve services addressed to the public.

To compensate the negative bias towards evaluation of innovation and salvage its role in improving quality, negative assessment of risk and the unpredictability involved in innovation must be balanced by its positive elements for the purposes of cooperation. The application of criteria to experimental initiatives should therefore be weighted and specific. Elsewhere, as far as the characteristics of innovation are concerned, the focus should be on initiatives or elements thereof that have indeed had an impact, albeit on a small-scale, and not on a negative assessment of mistakes as errors are also important in the learning process (Perrin 2002).

In some contexts, innovation can be a core aspect: when, for example, an initiative seeks to generate new forms of local public policies because changes in the context have made traditional approaches ineffective or when there is great uncertainty about which path to follow. In such cases, agents can consider the application of an emerging evaluation type, or developmental evaluation. This is an ongoing evaluation type that applies the evaluation's processes and logics to support initiatives and agents focused on social innovation in complex and unpredictable environments and systems (see Quinn Patton 2011 and Gamble 2008). In this case, the evaluator belongs to the team and accompanies the processes from design and planning, and encourages ongoing evaluation of the initiatives being undertaken in order to foster improvement in and adaptation of the initiatives and agents.

---

5 This tendency can be corroborated in the example of the five DAC criteria:
/ The agents of a cooperation initiative tend to encourage proven strategies and to reject the risk of assuming innovative solutions (relevance).
/ A traditional initiative features more accumulated information to formulate budgets and activities that use resources optimally in comparison with other initiatives that have not previously been proven (efficiency).
/ An initiative based on repeatedly applied intervention logic has more chance of achieving its results and targets (effectiveness).
/ The uncertainty associated with innovation makes experimental initiatives more likely not to achieve the envisaged impact or to generate undesired impact (impact).
/ Experimental initiatives involve more risks with regard to their demonstrability and sustainability when compared to others with strategies of intervention and exit validated through practice (sustainability).
3.3. Beyond criteria: flexible models focused on evaluation objectives and questions

The criteria establish general guidelines for reflection in order to answer evaluation questions. They are general dimensions of quality that must be specified flexibly for each evaluation. All or just a few of them may be applied or their significance weighted to suit the case. The criteria are neither exclusive nor exhaustive but rather complementary and their usefulness depends on the evaluation’s objectives and questions.

The guide does not therefore suggest that the criteria reviewed above should be used in a rigid, closed evaluation model. There is no best type of evaluation method and there are no generally applicable methodological recipes. The core element and the raison d’être of any model should be the objectives and questions behind the evaluation. What are the purposes of the evaluation? What questions should it answer? Depending on the evaluation questions, some criteria become more significant than others. Some will be particularly effective in systemising and organising the evaluation. The same is true for methodological tools.

The evaluation must be guided by its ends and not by methodologies. Only when there are clear objectives may the most suitable methodological approach be determined (DANIDA 2012: 16). The guide therefore opts for an evaluation model focused on the following categorisation of priorities:

- Objectives
- Questions
- Criteria
- Methodology

In accordance with this logic, criteria and methodological tools are defined and adopted to systemise and organise the questions and orient the search for information and analysis of data. If it is considered more useful and beneficial, an evaluation may also be structured in a different and alternative way to the criteria, and opt for different categories or concepts that allow for assessment of the questions raised. Indeed, ever-growing numbers of evaluations have not necessarily followed the logic of the DAC criteria as core elements of the model. They may be participatory evaluations that have no previously defined criterion or studies in which the criteria have been used implicitly, albeit not for structuring the evaluation analysis or report (see, for
example, Fujita 2010 and ODI 2008).6

The nature of the DC analysed in this guide, focused on the generation of local public policies, may also suggest orientation towards evaluation models that are more focused on processes than on results. The key elements for evaluation would therefore change from the products and results yielded by the initiative to the changes and improvements in the capacities of the different agents and populations involved. Assuming that the initiative is simply another influence on local public policies and services, evaluation would not so much seek categorically to attribute impact to the DC initiatives analysed, but rather to understand and accompany the general improvement processes in local policies that contribute to social cohesion.

In short, the most important thing is for each evaluation to have a specific design and methodology suited to the purposes and questions of evaluation. The core importance of the evaluation objectives and questions does not, however, imply that the evaluation team should not be critical. The evaluation also prompts “questions about questions” (W. K. Kellogg Foundation 2004). It is therefore important to analyse whether the questions reflect solely the needs and interests of donors (or other clients) and/or whether they assume exclusive ideological or political approaches. The evaluation team should also ask itself what perspectives and interests may be missing and try to incorporate key agents that have been excluded from the evaluation perspective. Concern for the diversity of agents and perspectives is reflected in trends that are increasingly highlighting the impact of evaluations on improvements in agents’ capacities. Evaluation is not limited to the provision of findings and recommendations applicable to the initiatives analysed, but should rather encourage participants to acquire new knowledge and skills in order to improve quality and prompt development in their institutions, systems, ideas and values (Quinn Patton 2008).

6 The application of a systemic approach has also been proposed to provide variables that are alternative and/or complementary to the criteria for orienting an evaluation (for further information see Fujita 2010; chap. 3).
4. Management of evaluation

This final chapter deals briefly with different phases of the evaluation management cycle and the core instruments of evaluation such as the terms of reference, the evaluation matrix and the final report.

The management tasks associated with evaluation can be divided into four basic phases: planning, design, performance and application, and widespread implementation of the results. An LG may establish a simple evaluation procedure on the basis of the different phases and needs thereof. As this procedure envisages the agents involved in the evaluation cycle and in each phase, the necessary activities, costs, established procedures, and expected products, etc., it would enable the LG systematically to take into account all the basic elements that must be guaranteed in order to perform an evaluation. An evaluation procedure that systemises phases would allow planning of the evaluation cycle to suit the specific characteristics of the LG operating structures.

The initial phase of the evaluation cycle –planning– establishes the needs and interests of the agents involved and therefore the evaluation’s purposes and objectives. The envisaged uses of the evaluation’s results and lessons will therefore also be considered. The LG must formulate the ends and the envisaged utility of the evaluation it is to undertake. This phase involves determination of basic elements such as the resources allocated to the evaluation or whether it should be internal or external.

A core element of the design phase of the evaluation is the formulation of the evaluation questions that, as mentioned previously, will determine the series of evaluation approaches and models to be applied. Here, the agents and perspectives involved in producing the questions are of crucial importance. In participatory evaluation, groups representative of the target population establish the questions. In other cases (internal, mixed, external), the agents who commission the evaluation set the questions and incorporate agents whose perspectives they prioritise. The more the target population take part in formulating the questions, the greater the evaluation’s legitimacy and transparency. Upon formulating the questions, it is also necessary to consider the utility of the evaluation and the application of its results (How can inclusion of the lessons learned in the work be guaranteed? Who will be responsible? What will the mechanisms be?).

In most cases, evaluations in cooperation are external (this is also true for the DC of LGs, which is recommendable because of the capacities required to evaluate impact in public policies) and involve the production of terms of reference (ToR) that establish the basic elements of the evaluation: questions, envisaged products, characteristics of the evaluation team, assigned resources, time available, etc. In formulating the questions, it is important to be as specific and as concise as possible. The quality of the conclusions and recommendations and their utility and applicability largely depend on the specificity and definition of the questions.
Lastly, in external evaluations the selection of the person or team in charge of evaluation must always be put out to tender.

The execution phase is basically the responsibility of the person or team that performs the evaluation. It involves the production of a plan of work (which envisages an execution schedule, a description of the proposed methodologies and implementation of the evaluation questions). The team’s work generally involves a stage of preliminary work (study of documentation and interviews, production of indicators and hypothesis), fieldwork (compilation of information in interviews, at workshops, and through participant observation, etc.), and work to analyse information and write reports (analysis of data and production of conclusions and recommendations).

Table 8: Example of reference term contents

| Background and justification of the evaluation | Basic information of the initiative/s to be evaluated: title; summarised description; logic of intervention: activities, results and expected objectives, etc.; donor and executor bodies; target population; time of execution; changes occurred during execution and current situation; available documentation; etc. Another core element is an explanation of the background and justification of the evaluation: ends and objectives of the evaluation, agents who are planning and commissioning the evaluation, etc. |
| Evaluation objectives and questions | Definition of the evaluation objectives. What accountability and learning objectives are established by the LG for the evaluation? How should evaluation allow for improvement in the quality of the DC initiatives and of LG public structures and policies? Likewise, evaluation questions that specify objectives will be itemised and the evaluation work will be focused on aspects prioritised by the LG. |
| Agents involved | Description of the different agents involved in the different phases of evaluation and their roles |
| Methodological considerations and requisites | Although the evaluation team is essentially responsible for the evaluation methodologies used, the LG may contribute approaches it considers fit in order to suit the evaluation to its needs. |
| Description of expected products | Description of the products expected from the evaluation that the evaluation team must submit to the LG: preliminary report, definitive report, executive summary, and its contents (basic index), styles of writing and extension. |
| Characteristics of the evaluation team | The nature of the evaluation team is defined in accordance with the objectives and characteristics of evaluation: number of members, prior knowledge and experience in evaluation, in DC, in the specific area of intervention, with regard to geographical context, etc. The selection mechanisms will also be defined. |
| Mechanisms of application and dissemination of results and recommendations | Description of the dissemination measures and widespread implementation of the expected products and of their conclusions, and the responsibilities of the evaluation team in this regard. |
| Schedule and work plan | Dates of execution of the evaluation and of the presentation of the expected products. |
| Available budget | Budget available for contracts. |
A key tool in the execution phase is the evaluation matrix that systemises the evaluation’s main contents: elements to be evaluated / evaluation questions, criteria or other dimensions with which the work can be structured, indicators, proposed methodologies, sources of information, etc. The matrix, which is developed by the evaluation team, is a tool with which to specify and refine the evaluation questions. The evaluators propose the content to the LG, which must assess whether the matrix is suited to and specifies the purposes of the evaluation and the general questions that orient it.

Upon completion of the execution phase, the draft evaluation report must be submitted so that the LG and the other agents involved can make comments, which the evaluation team can opt to incorporate or not to incorporate in the final version of the evaluation report.

Lastly, the final phase envisages application of the recommendations and widespread implementation of the results of the evaluation. This phase very much determines whether or not the evaluation fulfils its purpose of learning and improving quality. If the evaluation allows neither for improvement nor sharing knowledge on the lessons learned, then it is not relevant. The results of evaluations should be taken into account, for example, in future phases of planning of new DC initiatives. It may also be useful to establish mechanisms of citizen participation in order to monitor implementation of the recommendations: for example, by establishing a committee for monitoring and applying evaluation results that comprises representatives of the public and the initiative’s target groups. The lessons learned may be disseminated to a wider public through a broad range of mechanisms such as: publication of summaries, presentations and seminars, workshops for reflection or training, dissemination by Internet, databases of evaluation results, publications in journals, distribution to local agents, press conferences, etc. Communication strategies and products

<table>
<thead>
<tr>
<th>Quality dimension (criteria or other)</th>
<th>Evaluation questions</th>
<th>Indicators</th>
<th>Sources of information</th>
<th>Data compilation methodology</th>
<th>Data analysis methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main questions</td>
<td>Secondary questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

depend on the purposes of widespread implementation and on the nature of the agents to whom they are addressed.

LGs that wish to encourage the use of evaluation and a culture of evaluation in their structures must heed institutional elements and develop an evaluation policy that contemplates methodologies, budgets, roles and responsibilities, mechanisms for planning and applying results, etc., i.e. a strategy for improvement.

The specific way in which an evaluation policy is institutionalised may vary, although most multilateral cooperation agencies and bodies have independent evaluation departments or units that are accountable to policy managers (the cooperation system to have developed the greatest independence is the Swedish system, as the body responsible for evaluation, SADEV, is completely independent from the cooperation agency, SIDA). In the case of this guide, an evaluation unit could be established or a manager simply designated in accordance with the human and budget capacities of each LG.

The institutionalisation of evaluation processes may also be undertaken in accordance with the general objective of the DC to strengthen local public policies. Hence, evaluation must be understood as just another tool for learning in order to improve the quality of the initiatives and policies implemented, thus prompting better cycles of local policy planning and therefore achieving better impacts in regional social cohesion.
Methodology appendices

There follow two appendices on methodology. As they involve highly technical contents that would not necessarily be of interest to the technical and policy teams of the LG in charge of running evaluation processes, they have not been included in the main body of the guide. They should nonetheless be consulted both when commissioning external evaluations and performing internal evaluations.

For external evaluations (probably the most common), consultation of these appendices will enable LGs to understand the basic aspects associated with the choice of evaluation methodologies and instruments. It will also help to understand and assess the methodological proposals of teams of external evaluators. It can also enable LGs to help their external teams formulate methodological proposals in order to ensure that their interests and approaches are taken into account. Lastly, the contents featured here are useful for LGs in methodological monitoring of the evaluations that they commission.

Consultation of the appendices is essential for internal evaluations, albeit insufficient. It must therefore be complemented with the texts that appear in the bibliography or others that enable LG evaluation teams to articulate a suitable and thorough methodological proposal, which will ensure that the evaluation fulfils its accountability and quality improvement objectives.
Methodological tools are used in evaluation against a background of prolonged and persistent debate between quantitative and qualitative techniques. The positivist approach defends quantitative perspectives and tools: scientific techniques, hypothetico-deductive methodologies and statistical experimental or quasi-experimental models with which it is objectively possible to discover and explain the facts. The constructivist approach, on the other hand, tries to understand rather than explain, and opts for a qualitative perspective that analyses the multidimensional nature of phenomena (socio-economic, political, cultural, etc.) and focuses its attention on the diversity of perspectives and values that influence evaluations’ findings and conclusions.

Champions of experimental and quasi-experimental models argue that they represent the only way of attributing effects to an evaluated initiative. Non-experimental models (qualitative), they maintain, have no counterfactual (a control group showing the situation the target population would be in had the initiative not been run), and are thus unable to separate the effects caused by the initiative from those generated by external factors. As they are unable to rule out the influence of the context, they are unable to determine whether or not the effects studied are attributable to the evaluated initiative. Experimental or quasi-experimental evaluations of impact are therefore the only forms of rigorous and objective evaluations. This position is reflected by different agents of the cooperation system such as the previously mentioned NONIE, 3ie and the Center for Global Development (CGD), which foster initiatives to generalise rigorous experimental methods in the evaluation of cooperation initiatives (USAID 2009, CGD 2005).

Champions of qualitative approaches meanwhile respond that objectivity cannot be attained and that findings and conclusions depend on the perspectives adopted. They would ask: What is understood by rigour? Rigorous evaluations for whom and for what? From this point of view, experimental methodologies are not the only ones that can aspire to rigour and objectivity while counterfactual evaluation is a suitable way of analysing causality in fewer than a quarter of policy areas (Jones 2009). Qualitative methods therefore analyse complex, multidimensional processes and concentrate on changes in conducts and values of the agents involved, whereas experimental models run the risk of focusing on quantifiable and predictable effects, and leave aside analysis of...
unexpected impacts and changes in the environment that are essential in understanding development processes. Refusal to consider experimental models as superior and defence of the diversity of methodological approaches, in accordance with the needs of evaluations (even impact evaluations), is a position defended by agents such as the American Evaluation Association (AEA) and the European Evaluation Society (EES).

Notwithstanding theoretical debates, when dealing with an evaluation process it is important not to oppose methodologies, but rather to integrate and triangulate different tools for answering the evaluation's objectives and questions. A combination of quantitative and qualitative methods complements both types, enhances their advantages and limits their drawbacks. When choosing tools, it is also necessary to consider factors such as budget availability, time and information, the experience of the evaluation team, the geographical and cultural context of the initiative and the type of cooperation initiative (thematic, sector, etc.).

By way of conclusion and in line with the approach taken throughout the guide, we would opt here for a perspective that considers the best methodology to be the one most appropriate for each evaluation and its questions (Quinn Patton 2011). There are no methods that are superior to others, but rather methods more suited to specific evaluation contexts, objectives and questions. The skill of the evaluation team lies in identifying and applying the most suitable and appropriate methodologies.
APPENDIX 2. Description of some evaluation tools

Some evaluation tools are described below. The space available here does not allow for exhaustive description or a presentation of further examples. For more information, please consult the bibliography. When applying any tool it is important to envisage the inclusion of the specific perspective of women (and adapt it to their times and needs, prevent any obstacles to their participation, etc.).

### Table 11: Examples of evaluation tools

| Tool                        | Brief description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | When to apply it                                                                 | Advantages                                                                                       | Limitations                                                                                                                                                                                                 |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|----------------------------------------------------------|---------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Semi-structured dialogue    | The technique allows for information to be obtained through dialogues with people or groups. It attempts to overcome the limitations of questionnaires or surveys (closed subjects, non-inclusion of the themes proposed by the people participating, etc.) and generates exchange with the people who provide the information. To do so, the tool solely envisages the prior production of a list of subjects that should be dealt with in the evaluation (5-15 for example). Its application is flexible and designed to suit the exchange generated with the people or groups. Although during the dialogue the list helps to keep the focus on the basic subjects analysed in the evaluation, it should not be applied rigidly, but rather adapted to any unexpected changes and subjects arising in the exchange. It is necessary to respect people, listen carefully to the subjects they wish to deal with, even if they digress from the questions initially chosen. Selection of the people and groups to be interviewed is a core element. The choice will depend on the objectives and needs of the evaluation. Special care must be taken to avoid biases (selection of people who have more power, are more accessible geographically, are mainly men, who do not reflect the diversity of groups and agents, etc.). After the interview, it is necessary to analyse the data and assess its reliability and veracity using different criteria. The tool can be applied with a broad range of subjects depending on the evaluation needs: key informants (inaccordance with whether they are representative, technical knowledge, responsibilities, etc.), selected groups, families, etc. If performed with groups, it is useful to use tools with which to display the themes dealt with (a blackboard, posters, etc.) in order to facilitate dialogue and reflection. | It is a technique for general application. | It allows for information collected from other sources to be verified. | Dialogue may involve the risk of yielding solely opinions that cannot be compared with other sources. It requires the creation of a climate of trust. |
| Participant observation     | The participant observation technique originated with anthropologists who use it as a tool for understanding communities by spending long periods in their living systems. In the evaluation process, it is applied more simply and seeks to understand the processes studied through regular and ongoing participation in specific spaces or activities of the target population of the evaluated initiative. It allows for more effective comprehension than other methods of the perspectives of the participant people and groups with regard to the initiative’s results and impact. The spaces or activities for participation are selected in accordance with the objectives of the evaluation, with the information required, or with the hypotheses subject to comparison.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | This technique allows for qualitative enrichment of information collected using other tools. | It facilitates understanding of the perspectives and values of the participant people and groups. | It requires time.                                                                 |
The information and ideas arising from participant observation are systemised to suit the needs of the evaluation.

Outcome mapping (OM) is a methodology that was created in the year 2000 by the International Development Research Centre (IDRC) in order to document the outcomes of development initiatives by emphasising changes in conducts in the main agents involved. Outcome is defined as a specific type of result: modifications in the practices of people and groups. The methodology’s theoretical approach questions the central importance of attributing impact in the evaluation. It considers that attribution is a problem that is hard to solve given the convergence of agents and causes that explain the changes and because these often occur over a long period of time. It therefore focuses on analysing and evaluating the contributions of initiatives or organisations to the desired outcomes. Instead of evaluating the impact of a development initiative, it analyses changes in the practices of the agents involved in the action. Said approach makes learning a core element of monitoring and evaluation, and orients them to changes in people and groups and in their assumption of the initiatives as their own. The technique comprises three basic stages:

1. **Intentional design**
   - This stage involves different dynamics and workshops and allows for participatory definition of different elements regarding the initiative and its context: future prospects, the initiative’s mission, direct partners, the desired outcomes, signs of progress, the map of strategies and organisational practices.

2. **Outcome and performance monitoring**
   - This stage features the participatory production of monitoring priorities and of its main tools: outcome, strategy and performance journals. The tools reflect the three components of intervention upon which OM is based: changes in practices, intervention strategies for promoting said changes, and the initiative’s organisational operation.

3. **Evaluation planning**
   - The last stage defines the evaluation plan. There is insufficient space in the guide to deal with use of methodology. For further information:

The most significant change (MSC) technique is a form of participatory monitoring and evaluation that was developed by Rick Davies in 1996. The purpose of the process is to identify the most relevant story of the change yielded by an initiative and/or an organisation. To do so, the main steps of the technique are:

- Evaluations geared to learning.
- Large-scale initiatives.
- It is more of an approach to evaluation than a specific tool.
- Its application shapes and orients the whole evaluation design.

It allows for the identification and analysis of changes in the different agents involved.

It allows for assessment of the adaptation capacities of initiatives.

It does not attribute results and impacts to initiatives.
Most significant change (MSC)

Compilation of significant change (SC) stories. The stories of an initiative's target population are compiled on the basis of a simple question to the participant people or groups such as: “Since last month, what do you think the most significant change in the quality of life of the people of this community has been?” SC stories can be collected using different tools (surveys, interviews, discussion groups, direct drafting by the people or by the initiative's staff, etc.). The basic information of each SC is compiled in a format of one page or less, which also features record of who has compiled the story and why the person thinks it matters. SC stories may be ordered by change domains (between 3 and 5), which are broad categories to help produce, classify and process them.

Selection of the most important SC stories
Groups of local agents discuss the SC stories and present those they consider most significant to a higher level within the organisational structure. This, in turn, selects the most important SC stories and conveys them to a higher level. This outline is repeated until the most significant change (MSC) of all the stories is selected. Use may be made of the structure of the organisations in the selection process or a specific structure may be established to select the SC stories. There are several methods of selection (voting, consensus, scoring, etc.). The reasons for selection must be documented.

Feedback and verification
The results of the process are shared with the participant groups and other agents, with an explanation of which SC have been selected and the reasons why. Verification is also performed to ensure that the MSC is a suitable reflection of reality.

Manual available at: www.mande.co.uk/docs/MSCGuide.pdf

Network diagrams
The network diagrams technique allows for analysis, with different participants, of relations and contacts among agents and of the impacts of the initiative thereupon. On the basis of work with different groups, diagrams are produced that show relations among agents. For example, evaluation of the impact of an initiative on the coordination mechanisms among different departments of a local government involves work with each to produce a diagram in which the main departments shown on a card or on a blackboard are joined with arrows that characterise the different types of relations. Arrows of different shapes or colours reflect relations in accordance with the criteria that require analysis such as: decision flow, thematic coordination, shared responsibilities, nature or quality of relations, frequency of contacts, etc. In addition to the departments, this exercise can be repeated with specific groups whose perspective needs to be incorporated (women, subgroups formed according to decision-making capacity and influence, participants and non-participants in the initiative, etc.). The diagrams allow for evaluation of impact, and of differences in perceptions of changes among different agents.

Evaluations focused on learning.
Complex initiatives with different and emerging results, unexpected changes, with large organisational structures.
It is more of an approach to evaluation than a specific tool. Its application shapes and orients the whole evaluation design.

It facilitates the identification of unexpected changes.
It facilitates participation, is suited to different cultural contexts and educational levels, and does not require technical knowledge of participants. It can provide plentiful material to back up evaluation analyses.

It requires time.
It is biased towards successful stories.
The MSC selection process is subjective and hierarchical.
Spaces for the return of results must always be provided. The following image is a very schematic example (in which each circle represents an agent and each line a type of relation) of the resulting diagram type:


ALNAP guide for humanitarian agencies, ALNAP & IECAH, Madrid.

ASOCAM (2009): Monitoreo y evaluación de acciones de desarrollo orientadas al impacto, ASOCAM, Quito.


CLAMA (2009): Mapeando Alcances. Un manual práctico para el uso de Mapeo de Alcances en procesos de Desarrollo en Comunidad, Latin American Centre for Outcome Mapping, CLAMA.


DFID (2009): Building the evidence to reduce poverty. The UK’s policy on evaluation for international development, UK Department for International Development (DFID), London.


GTZ (2007): Capacity WORKS. El modelo de gestión de la GTZ para el desarrollo sostenible, GTZ, Eschborn.

GUANZIROLI, C. et al. (2007): Metodología de evaluación del impacto y de los resultados de los proyectos de cooperación técnica, Inter-American Institute for Cooperation on Agriculture (IICA), Brasilia.


LARRÚ, J. M. (2007): «Impact Assessment and Evaluation: What it is it, how can it be measured and what it is adding to the development of international co-operation», MPRA, No. 6.928.


UCLG (2009): UCLG position paper on aid effectiveness and local government, UCLG.

UNCETA, K. et al. (s.f.): «Objetivos del milenio, financiación del desarrollo, y eficacia de la ayuda 2000-2010: necesidad de un análisis integrado y de un enfoque alternativo», UPV/EHU, XIII World Economy Meeting.


Various Authors (2011): Proceso consultivo sobre la eficacia de la ayuda a nivel local.


WORLD BANK (2004), Monitoring and evaluation: some tools, methods and approaches, World Bank, Washington D.C.

Some websites upon which to find information on evaluation

Development Co-operation Directorate (DCD) – OECD
www.oecd.org/dac/evaluation

United Nations Evaluation Group (UNEG)
www.uneval.org
Evaluation Website – EU External Relations
ec.europa.eu/europeaid/how/evaluation/

Independent Evaluation Group (IEG) – World Bank
www.worldbank.org/oed/

Network of Networks for Impact Evaluation
www.worldbank.org/ieg/nonie

International Initiative for Impact Evaluation
www.3ieimpact.org

United Nations Development Programme (UNPD)
web.undp.org/evaluation/index.html

Plataforma Regional de Desarrollo de Capacidades en Evaluación y Sistematización de América Latina y el Caribe
www.preval.org

International Development Research Center
www.idrc.ca

Foundation for Advanced Studies on International Development (FASID)
www.fasid.or.jp/english/

Plataforma Latinoamericana de Gestión de Conocimientos
www.monitoreoyevaluacion.info

Institute of Development Studies
www.eldis.org

Center for Global Development
www.cgdev.org/

European Evaluation Society
www.europeanevaluation.org/

American Evaluation Association
www.eval.org/

Red de seguimiento, evaluación y sistematización en América Latina y el Caribe (ReLAC)
www.relacweb.org/

Outcome Mapping Learning Community
www.outcomemapping.ca/

Overseas Development Institute (ODI)
www.odi.org.uk

Swedish International Development Cooperation Agency
www.sida.se/English

Canadian International Development Agency
www.acdi-cida.gc.ca/

Agencia Española de Cooperación Internacional para el Desarrollo (AECID)
www.aecid.es

Agence Française de Développement
www.afd.fr

Danish International Development Agency
www.um.dk/en

Japan International Cooperation Agency
www.jica.go.jp/english/
Swiss Agency for Development and Cooperation  
www.sdc.admin.ch/En caché - Similares

Red Latinoamericana y Caribe de Evaluación y Monitoreo  
redlacme.org/

Plataforma Latinoamericana de Gestión de Conocimientos para el Desarrollo Local en Áreas Rurales  
www.asocam.org

Red de Monitoreo y Evaluación de Política Pública en Colombia  
redcolme.ning.com/

Rede Brasileira de Monitoramento e Avaliação  
redebrasileirademea.ning.com/

UNICEF  
www.unicef.org/spanish/evaluation/

The Evaluation Center - Western Michigan University  
www.wmich.edu/evalctr/

Program Development and Evaluation – University of Wisconsin  
www.uwex.edu/ces/pdande/evaluation/

Monitoring and Evaluation NEWS  
mande.co.uk

Community Sustainability Engagement Evaluation Toolbox  
evaluationtoolbox.net.au/

Resources for Program Evaluation  
gsociology.icaap.org/methods

Evaluation Capacity Development Group  
www.ecdg.net

Impact Alliance  
www.impactalliance.org

The Active Learning Network for Accountability and Performance in Humanitarian Action  
www.alnap.org

Centre Européen d’Expertise et d’Évaluation  
www.eureval-c3e.fr

Participatory Planning Monitoring & Evaluation  
portals.wi.wur.nl/ppme/content.php

Evaluateca  
evaluateca.wordpress.com/

Debate on development  
www.portal-dbts.org/31_evaluacion_cast.html
URB-AL III is a decentralised regional cooperation programme run by the European Commission, the aim of which is to help increase the level of social cohesion in sub-national and regional communities in Latin America.

Headed by Diputació de Barcelona, the URB-AL III Programme Orientation and Coordination Office supports the implementation of the programme and provides technical assistance and accompaniment to different projects with a view to realising their objectives.