GOBIERNO DE ESPAÑA
MINISTERIO DE POLÍTICA TERRITORIAL Y FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE POLÍTICA TERRITORIAL Y FUNCIÓN PÚBLICA
INSTITUTO PARA LA EVALUACIÓN DE POLÍTICAS PÚBLICAS

GUIDELINES FOR THE IMPLEMENTATION EVALUATION OF PUBLIC POLICIES

INSTITUTE FOR THE EVALUATION OF PUBLIC POLICIES

2020

GOBIERNO DE ESPAÑA
MINISTERIO DE POLÍTICA TERRITORIAL Y FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE POLÍTICA TERRITORIAL Y FUNCIÓN PÚBLICA
INSTITUTO PARA LA EVALUACIÓN DE POLÍTICAS PÚBLICAS

AUTHORS:

-Celia Casillas

-Laura García

-Elena Lago

- Mónica Macía

Director: Celia Casillas

Guidelines for the implementation evaluation of public policies

Secretariat of State for Territorial Policy and the Civil Service

Publisher: Ministry of Territorial Policy and the Civil Service 2020

General catalogue of official publications:
http://publicacionesoficiales.boe.es

NIPO: 277190534

Contents

INTRODUCTION

Public policy evaluation is considered a tool for improvement and learning of public policies, and for accountability in government action. Within the different approaches to evaluation, comprehensive assessment considers evaluation to be a process that combines the political-strategic analysis of public action with the analysis of its operational aspects. It also takes into consideration the entire life cycle of public policy. This is because the results of public policies cannot be separated from the consequences that deficiencies in their design or implementation may have on said policies. Following Rossi, Freeman and Lipsey, "evaluation is no more than one of the ingredients of an inherently political process (Maldonado Trujillo and Pérez Yarahuán, 2015).

The comprehensive assessment approach requires us to follow an evaluation methodology that encompasses all the stages of public action: design, implementation, and results and impacts. Since 2007, the National Agency for Evaluation and Quality (AEVAL in Spanish) has applied this comprehensive approach to its evaluations and has developed its own methodology which is described in several documents and especially in its practical guide for evaluation design and execution with an AEVAL Approach, 2015.

After the dismantling of AEVAL, the Institute for the Evaluation of Public Policies as a body of the General State Administration Services to promote the culture of evaluation of public policies and the formulation and dissemination of methodologies for evaluation[1], has sought to facilitate the use of tools that help to assess any policy plan or programme and contribute to the institutionalisation of evaluation, its integration in public administration from the planning stage onwards. These tools include specific methodological guides on the different dimensions of a comprehensive assessment of public policies, aimed both at evaluators and the managers or public officials in charge of commissioning said evaluations.

This series consists of the following guides, in addition to this Guide for Evaluating Public Policy Implementation: Guide for the Evaluability Assessment of Public Interventions, Guide for Evaluating Public Policy Design, and Guide for Evaluating Public Policy Results, all published in 2020 by the Institute for the Evaluation of Public Policies[2].

Public policies are formulated in order to solve problems or requirements that improve the quality of life in society. To achieve this, certain objectives are established and their fulfilment is

[1] Royal Decree 307/2020 of 11 February which outlines the basic structure of the Ministry of Territorial Policy and the Civil Service. Article 2.5.

[2] All the guides are published on the website of the Ministry of Territorial Policy and the Civil Service in the Institute section:
https://www.mptfp.gob.es/portal/funcionpublica/evaluacion-politicas-publicas/Guiasevaluacion.html

measured by the results and impacts of said public policies. When the results and impacts are not the awaited ones, it is necessary to identify whether this is due to errors in the formulation or due to defects in the design and implementation of the public policies. That is to say, in the strategies, activities or measures that are defined for its deployment; in the organisations responsible for its management and execution; or in the resources allocated for the intervention. These aspects are dealt with in implementation evaluation, where the goal is to assess the progress or development of a public intervention and its deployment tools, identifying the risks and factors that may influence obtaining the results and fulfilling the defined objectives, and proposing recommendations for the improvement of the intervention.

Taking into consideration the above, the goal of this Guide for Evaluating Public Policy Implementation is to offer some simple guidelines for assessing the implementation of public interventions with two different target groups. On one hand, we have the persons or bodies in charge and the managers of the interventions who may thus have a general overview of the dimensions and contents of the implementation evaluation. On the other hand, we have the evaluators, following the implementation evaluation process, with examples and techniques.

The first part of this document, “General Questions” provides a simple response to basic questions on implementation evaluation in a question and answer format: what it is, what does it consist of, why is it recommended, and how to perform an implementation evaluation.

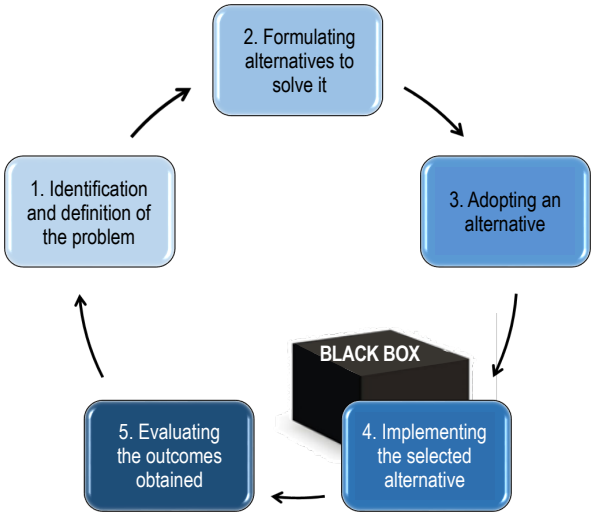
The second part “Methodology of Implementation Evaluation” enters into the details of the analyses required to perform the evaluation: the analysis of the deployment instruments and analysis of the level of implementation. Finally, it examines the evaluation criteria and questions as well as the evaluation techniques that may be used.

PART ONE. GENERAL QUESTIONS

What is implementation evaluation?

Public policies may be defined, according to Tamayo (1997) as the "set of objectives, decisions and actions undertaken by a government to solve the problems that are deemed a priority by the citizens and the government itself at a certain point in time". The sequential approach of public policies suggested by Lasswell (1962) breaks down the life cycle of public action into five stages: Identification of the problem, formulating alternatives to solve it, adopting a concrete alternative, implementing the alternative, and evaluating the results.

Figure 1. Design evaluation within the hierarchy of evaluation. Source: Author's own based on Rossi, Lipsey and Freeman (2003).



The implementation stage[3] therefore consists of setting in motion the operational plans, the administrative actions (procedures), the provisions, services, or instruments for economic or non-economic promotion, the investments, as well as the resources available for implementing the public policy (AEVAL, 2015).

Some authors define implementation as ***"the 'black box' of the intervention"***, which holds "all that is done and takes place in the programme: actions, decisions, resources, persons and interactions. It is expected that the combination of all these factors will produce the awaited results. These constitute the programmes (resources, persons, interactions, actions, etc.), as conceived of as something external as they are 'outside' the limits of the intervention (Ligero, 2016: 77). the results are merely the outputs which may even be conceived of as something external as they are 'outside' the limits of the intervention (Ligero, 2016: 77).

[3] The Spanish Royal Academy of Language defines implementation as the "action and effect of implementing" and to implement as "to set into motion or to apply methods, measures, etc. to execute something". (RAE, 2018).

The evaluation which is the main goal of this stage of public action is known as implementation evaluation or process evaluation (this Guide opts for the first term). Likewise, there are various definitions of this type of evaluation, among them that of the OECD, which defines it as that which "focuses on the internal dynamics of the implementing organisations, their policy instruments, their mechanisms for providing services, where the management procedures and the links between these components".

For the purposes of this Guide, implementation evaluation is defined as

> ***the evaluation that performs a systematic analysis of the operational dimension of the public action through its deployment and launch by means of activities and measures, as well as the analysis of the internal mechanisms (functioning and internal organisation) and the resources earmarked by the organisations in charge of executing it, in order to make an evidence-based assessment of said deployment and execution.***

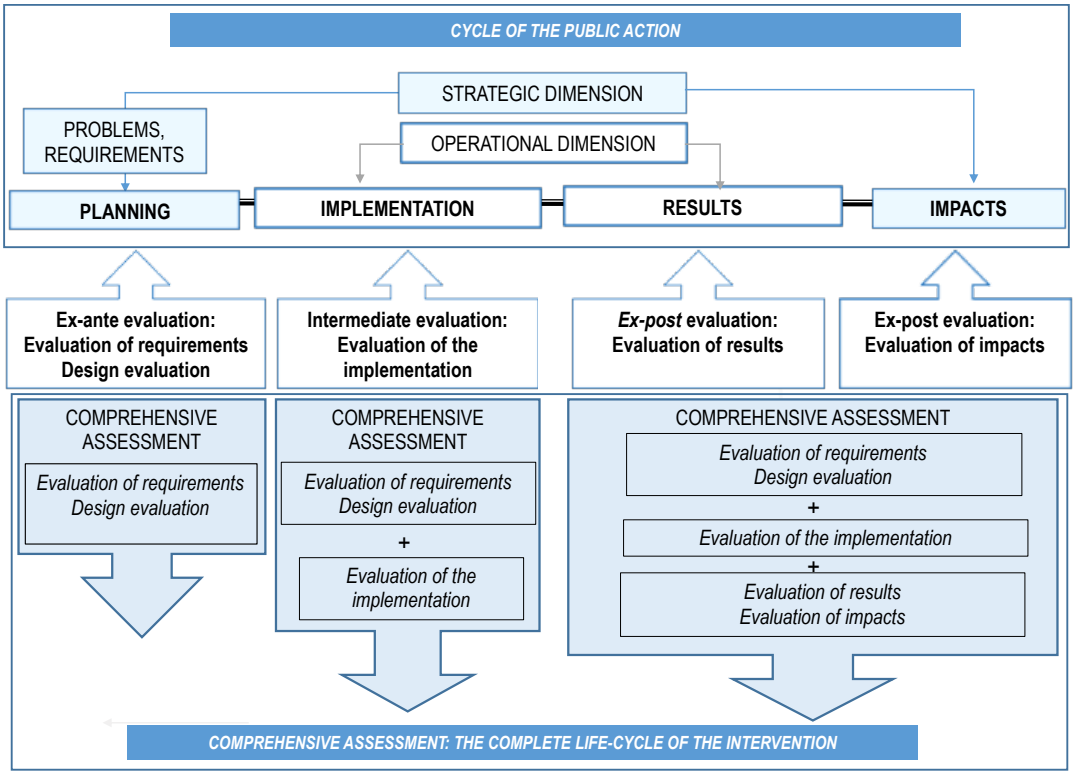
Implementation

Comprehensive assessment[4] considers that public policies are action processes characterised by their complexity and inter-relatedness with public problems, which require an approach that includes the problems as well as their solutions as opposed to the fragmented and decontextualised analyses offered by classical evaluation.

This approach is highlighted because it integrates the strategic and political analysis of policies (including problems, assessment, formulation and formalisation of the intervention and impacts) with the analysis of the more operational aspects of its deployment (resources, processes, activities, outputs, results). It is within this second analysis that the implementation analysis or evaluation is framed, thus making it a stage inside the process of executing a comprehensive assessment. The success or failure of a public intervention may be due to an error in its conceptualisation, to deviations in its implementation, or to errors in performing the activities required to bring about the desired changes to solve the detected problem. For all of these reasons, implementation evaluation is an essential part of comprehensive assessment which often allows us to explain the results of an intervention.

[4] Comprehensive assessment is "the systematic and reasonable process for knowledge generation, by compiling, analysing and interpreting information, aimed at the global comprehension of a public intervention –be it a policy, plan, programme or standard– to achieve an assessment based on evidence, considering its design, implementation and effects (results and impacts)" (AEVAL, 2015) .

Figure 2. Cycle of public action and comprehensive assessment. Source: Author's own based on AEVAL (2015).



Due to its methodological and didactic aims, this Guide focuses solely on implementation evaluation, assuming that the evaluator has already taken into consideration the analysis of the intervention formulation and design within the complete process of a comprehensive assessment. The methodology for performing the design analysis is described in detail in the Guide for Evaluating Public Policy Design, drafted by the Institute for the Evaluation of Public Policies in 2020.

Many of the pointers offered in this Guide are common to the general theoretical corpus on public policy evaluation, but others are the result of the expertise[5] of the evaluators of the Institute for the Evaluation of Public Policies.

What is the role of implementation evaluation?

The usefulness of this type of evaluation depends upon the current stage of the life cycle of the intervention to be evaluated and the degree of execution of the instruments and activities for its deployment: In any case, it is recommended that the implementation evaluation be

[5] Following María Bustelo (in Maldonado Trujillo and Pérez Yarahuán, 2015: 11), "the theory of evaluation has been built largely on the basis of inductive rather than deductive reasoning, from experience and including practical attempts to solve social problems in the theory. This gives it a theoretical status (...) in constant interaction with praxis".

performed whenever it is detected that the public intervention has not been executed as intended by the planners, an intermediate evaluation being the most likely means. In other words, when the execution and enforcement deadline of the intervention was not reached.

In an ***ex-ante evaluation***, when the intervention has not commenced its implementation or execution, it can help to improve the intervention design and is therefore considered a formative evaluation. It is performed in order to correctly define the deployment instruments, mainly those that refer to the information systems that feed the mechanisms for monitoring and evaluation (monitoring and results indicators), the organisational instruments and the mechanisms for coordination, among others. It makes the intervention evaluable after its completion, making it ready for an ex-post evaluation.

In the case of interventions ***at an intermediate evaluation stage*** and during the execution of the intervention, implementation evaluation may be useful for the early detection of problems and errors in executing the deployment activities or measures, and in the intermediate results that may influence the achievement of the initial objectives. It helps us to learn about the progress of the intervention, identify critical areas, take corrective measures, and make recommendations to re-direct the programme if necessary. Its objective is mainly to provide knowledge and improve the intervention, and therefore may be deemed, as in the previous case, a formative evaluation. It is especially useful in interventions that deal with dynamic or complex problems that may undergo modifications to their definition, components, causes, or effects owing to changes in some of the related structures, problems, context or components. These modifications may have several causes. Rein and Schon (2016) explain that, as a matter of fact, “many policy changes consist solely of adapting the policy to changing situations. Thus, the cumulative effects of many of these adaptations may entail a re-framing”.

When performed on a completed programme or policy ***(ex-post evaluation),*** implementation evaluation lets us understand what has taken place between the objectives and the observed results and impacts if these are not the awaited ones, as well as identify the factors responsible for achieving these results. It will be possible to judge if certain mediocre or deficient results are due to a design error in the activities or measures, in the implementation deficits or problems, or in the response of the target population to the programme, or to any other factor. It is considered ***a summative evaluation,*** as it allows us to draw conclusions regarding critical aspects that condition the results.

What does implementation evaluation consist of?

Implementation evaluation analyses and assesses the elements that are a part of the ***logic of the intervention***, or the internal line of argument of the intervention, consisting of the formulation and design (sequentially structured into objectives, activities and resources, the underlying causal hypotheses, and the generated outputs).

The logic of the intervention is also examined in the design evaluations, as the line between the two evaluations are blurred with regard to this point, because there are elements that may be deemed common to both, although the analyses performed in each evaluation may differ in their approaches, objectives and depth. The difference lies in the fact that design evaluation focuses on the relationship between the intervention and the problem, and on causal theory, whereas implementation evaluation focuses on the internal components of the design and on the existing relationship between the design, the outputs and the results and impacts, as may be observed in the figure below.

Figure 3. Chain of causal relationships. Source: Author's own.

The methodology of implementation evaluation considers the following components:

1.- Analysis to assess the implementation of an intervention

From a didactic perspective, the analyses that seek to assess the implementation of an intervention are divided into:

1. *Analysis of the deployment:* All public interventions are deployed as instruments or components that put the established strategies into operation in order to achieve the objectives defined in the planning stage. This analysis focuses on the assessment of said instruments or components of deployment, constituting of activities (within a wide concept of the term which includes measures, actions, plans, etc.) that are implemented in organisations through processes, with resources earmarked for them (called inputs) and which in turn produce outputs, the latter being assets, services or immediate transformations that are obtained via the implementation or execution of the activities. They constitute the first level of results of an intervention and are necessary to obtain the awaited effects or impacts (outcomes). The analysis assesses the appropriateness, availability, sufficiency, and proportionality of each element, as well as its effects on the obtained results and the awaited impacts. It also analyses the organisational tools, the key stakeholders of the intervention, their roles and the mechanisms for participation, coordination, etc. The information systems and mechanisms for monitoring and evaluation defined in the planning are also elements to be analysed in the deployment.

2. *Analysis of the level of implementation*: This analysis approaches the level of implementation of the intervention at the moment of the evaluation, reviewing the fulfilment of the intermediate targets, the projections regarding the execution, the assessment of the criticality of the measures, the risks associated with each stage and the consequences of the deficiencies and deviations of the implementation for obtaining the results and to achieve the objectives.

2.- Evaluation criteria and matrix

The ***criteria*** provide benchmarks (yardsticks, standards, principles, etc.) to obtain useful information in order to assess the evaluation. Those applicable to the implementation evaluation that may be taken as reference are: suitability, coverage, coherence, complementarity, coordination, effectiveness, efficiency, implementation, participation, proportionality, sustainability, and transparency. Those that are required may be used, that is to say, only some or all, and ad hoc criteria may even be created according to the requirements of the evaluation.

The ***questions*** are the basic unit of research, which may be defined as the queries and hypothesis to be confirmed which allow us to execute the evaluation. The list of questions and their associated criteria, as well as the sources of information, measurement indicators, techniques and tools are included in the ***evaluation matrix***, which constitutes the tool that logically integrates all these elements, containing the focus and scope of the evaluation.

Figure 4. Evaluation criteria in implementation evaluation. Source: Author's own based on AEVAL (2015).

Criteria associated with implementation evaluation	
Suitability	Degree of consensus among all stakeholders, especially the managers and the target group, on the development of the intervention and its instruments for coordination and participation.
Implementation	The manner in which a public intervention has been implemented or set in motion.
Coverage	The extent to which a public intervention reaches a segment or all of the target population.
Internal coherence	Consistency between the designed and implemented measures and the outputs and results.
Complementarity	Degree of alignment and *ad intra* interactions, that is, between the measures of the intervention, detecting synergies or antagonisms, and *ad extra* interactions, between these measures and other measures or policies.
Coordination	Degree of adjustment of the relationships between the managers, units, and institutions that are involved in the deployment.
Participation	The degree and manner of participation of the stakeholders during the implementation of the intervention.
Effectiveness	The extent to which the deployment activities or measures of the intervention obtain intermediate outputs or results in line with the planned objectives, within the established deadlines.
Participation	The degree and manner of involvement of the stakeholders in the intervention.
Proportionality	The extent to which the implemented measures are proportionate to the (organisational, resource-based, etc.) efforts to achieve the sought results.
Sustainability	Assess whether the measures set in motion and their outputs can be maintained over time.
Transparency	Degree of stakeholder knowledge and perception regarding the adopted measures, as well as the decision-making process.

3.- Analysis techniques in implementation evaluation

This triangular approach is a requirement of comprehensive assessment, as it considers all the theoretical-scientific perspectives that are deemed relevant and useful for evaluation. For this, all types of ***techniques and tools*** are used. This Guide mentions the most commonly used ones.

How is an implementation evaluation performed?

Implementation evaluation, as mentioned earlier, is an evaluation process focusing especially on the analysis and assessment of the launching or deployment of public action. The evaluation unfolds on the initiative of the manager or person or body in charge of the intervention, whose decision is usually reflected in an initial document that contains the analysis of the commission, and is developed by means of the analyses mentioned in the previous section.

The duration of the evaluation will depend on the complexity of the intervention and its nature, characteristics, and conditions, which include the resources allocated for the evaluation.

The process concludes with an evaluation report, which must describe the result of the investigation, the different analyses performed and the findings obtained, usually following a structure based on the evaluation queries used and their associated criteria. The final report must include a conclusions and recommendations section, always based on the obtained evidence. If we think of evaluation as another public intervention, then we may close the cycle with a follow-up of the evaluation. The recommendations of the aforementioned AEVAL Guide 2015 may be followed when drawing up the report, as well as for the monitoring activities of the evaluation.

PART TWO. METHODOLOGY OF IMPLEMENTATION EVALUATION

There is a significant amount of consensus in the literature on the need for implementation evaluation when it is detected that the public intervention is not being undertaken as envisioned by the planners. The success or failure of a public intervention may either be due to an error in its conceptualisation, to deviations in its implementation, or to errors in performing the required activities for achieving the desired changes to solve the detected problem.

The traditional concept of evaluation has focused on the results of public interventions based on a theory and hypothesis of intervention, and certain clearly stated objectives; the presence of a single stakeholder that takes logical and unequivocal decisions freely; and finally, a concept of the policy cycle as a linear succession of its different stages.

However, what's certain is that "the formulation and implementation of public policies is characterised by their bounded rationality derived from the different interpretations of the stakeholders of the public problem that gives birth to the policy; the contingencies of the organisations whose features (structures, processes, resources, etc.) may affect the achievement of the foreseen objectives, and finally, the inability of the government as sole stakeholder to solve public problems" (Ruiz, 2013).

To be noted

It is important to distinguish between the terms "implementation stage of an intervention" and "implementation evaluation".

The implementation stage of an intervention *begins with the adoption of a policy, plan, or programme, until the evaluation of the results, if we follow the life cycle of public action drawn up by Tamayo (1997).*

Implementation evaluation *refers to the evaluation process that sheds light on how the implementation is performed and the outputs that are obtained with the goal of achieving the objectives proposed by the intervention.*

This gives birth to the reasons for performing an ***implementation evaluation***, based on the lack of development of the expected actions, the lack of information regarding the intervention, and the need to learn about the processes or assistance to interpret the results (Weiss, 2016). Because "to neglect the comprehension of implementation leads to vagueness when determining, in many evaluation investigations, if the implementation programme or system or both, are responsible for the demonstrated errors in obtaining results (Huey-Tsyh Chen and Peter H. Rossi 2016)[6].

Implementation evaluation is that which focuses on the analysis of the operational dimension of public action, that is to say, on the deployment and launch of activities and measures in interventions, as well as the analysis of the internal mechanisms (of functioning and internal organisation), and of the resources earmarked by the organisations in charge of executing them. In the literature on evaluation, some authors call this type of evaluation "process evaluation". Although both are accurate, this Guide has opted for the term "implementation evaluation".

[6] Chen and Rossi (2016) propose an implementation system that acts as an "organisational arrangement".

The evaluation methodology presented in this section includes the specific analyses intended to assess the implementation of an intervention[7] according to the following scheme:

1. *Analysis of the deployment:* This section deals with the analysis of the programming and organisational instruments of the deployment, the relevant stakeholders of the intervention, the resources allocated for the intervention, the mechanisms for monitoring, evaluation, and accountability, and finally, the instruments for communication and dissemination.

2. *Analysis of the level of implementation:* It deals with the level of implementation, aided by the review of the intermediate goals, the execution outlook, the assessment of the criticality of the measures, the risks, and different tools.

Next, it describes the *criteria* that are generally included in an implementation evaluation, as well as an *evaluation matrix* proposal. Finally, in the last section, the most-used *techniques and tools* in evaluation are shown.

It is important to highlight that similar to design or results evaluation, implementation evaluation is an essential part of comprehensive assessment, and is therefore necessary to conclusively assess the success of a public intervention. Nevertheless, due to its methodological and didactic aims, this Guide focuses solely on implementation evaluation, assuming that the evaluator has already taken into consideration the analysis of the intervention formulation and design within the complete process of a comprehensive assessment. The methodology for performing this design analysis is described in detail in the Guide for Evaluating Public Policy Design, (Institute for the Evaluation of Public Policies in 2020)[8].

[7] For more details on this type of analytical structure, see, among others, María Velasco (2007) Distintos instrumentos para un mismo fin. Los instrumentos de las políticas públicas como herramienta para el análisis, VIII Congreso Español de Ciencia Política y de la Administración, 18-20 September 2007, Valencia (Unpublished).
[8] Guide for Evaluating Public Policy Design (Institute for the Evaluation of Public Policies, 2020). The components of design evaluation are:
- Analysis of the intervention formulation: analysis of the public problem, of the choice of public intervention (including the analysis of the alternatives) and the context of the intervention.
- Analysis of the intervention design: analysis of the theory of action and the theory of change or the logical-causal relationship of an intervention.

1. ANALYSIS OF THE DEPLOYMENT

All public interventions are deployed as instruments or elements that operationalise the execution strategies in order to achieve the objectives defined in their design. These components, along with the structuring of the objectives (strategic, specific, and operational) and the results and impacts, constitute the theory of action or implementation within the logic of the intervention .

The instruments or components of the deployment consist of different activities (in a wide understanding of the term which includes measures, actions, plans, etc.) that are implemented in organisations through processes (implementation chain) with resources earmarked for them (called inputs), and which in turn produce outputs, the latter understood as assets, services or immediate transformations that are obtained via the implementation or execution of the activities. They constitute the first level of results of a specific intervention and are necessary to obtain the awaited effects or impacts (outcomes).

Figure 5. The results chain of a public intervention. Source: Author's own.

With reference to the results chain as an expression of the sequential logic of the intervention, it may be stated that implementation evaluation addresses the inputs, the activities, the outputs, and the applicable or non-applicable connections between them. It attempts to analyse how each link in the chain functions and how they affect the relationships at each level and what they seek to achieve in pursuit of the awaited results and impacts. With regard to the inputs, it examines their suitability and availability for each activity. With regard to the activities, it analyses their suitability, proportionality, and capacity to generate the outputs, and whether these outputs lead to the desired results or not.

In order to perform this analysis the evaluator must identify and validate the underlying hypotheses at different levels of the chain in order to make an assessment based on criteria such as internal coherence, coordination, complementarity, and the degree of implementation of the evaluated intervention.

[9] The logic of the intervention is created by the most operational part of the public action, the theory of the action, and the most strategic part, the theory of change or causal theory of the intervention, which refers to the manner of producing the changes required to fulfil the objective of an intervention or a desired final change.
[10] The implementation chain consists of all those processes that are directly or indirectly involved in the action to implement a public programme or policy.

We must remember that this sequence of results may be more or less explicit when performing an implementation evaluation. If there are gaps in the evidence of every link in the chain, the evaluator must reconstruct the chain and the underlying hypotheses at each level on the basis of documentary revision, interviews with the persons or bodies in charge, and the participation of the stakeholders. Even when the results chain is explicit, the evaluator must equally compare the hypothesis of each level to ensure its validity.

In any case, stakeholder participation at this stage of reconstruction or validation of the results chain allows us to improve aspects related to the appropriateness and validity of the design, the adaptation and commitment of the stakeholders to the intervention, and the coordination and cooperation between them.

Finally, it must be considered that public policies in general are implemented through public organisations and public services whose administration, organisational management and resources are elements to be studied in the implementation evaluation, given that their deficiencies or inefficiencies may affect the achievement of the results.

Below is the detailed analysis of all these components.

1.1. Programming instruments for deployment

The first step of this analysis is to identify and ***describe the deployment activities of the intervention, that is to say, the measures, sub-measures, actions, activities, and outputs of the intervention***[10]. The goal is to obtain an orderly view of the planned sequence of development of the intervention and how the activities and projects are to be performed.

Depending on the nature of the intervention to be evaluated (a policy, a strategy, a plan, or programme), the activities may be more or less extensive and complex and may be deemed interventions in and of themselves and therefore, open to evaluation. If the intervention to be evaluated is a public policy, as in the case of the AEVAL evaluation of administrative burdens in company creation, the evaluator shall encounter a series of deployment activities that individually constitute an intervention (they may be strategic plans, action plans or concrete measures and actions), which makes the implementation evaluation of said public policy a complex affair.

The intervention may be defined as a Plan that initially appears to be a concrete intervention but is actually a set of different interventions within the same framework.

[11] There may be different denominations in the plans and programmes: measures, initiatives, projects, sub-measures, actions, activities, etc.

Implementation evaluation must approach each measure and their implementation individually, as well as jointly, especially when dealing with the underlying logic of each measure and the general logic of the intervention (plan, programme, strategy...), which markedly increases the complexity of the evaluation.

On the other hand, the evaluation will be less complex if the intervention to be evaluated is a sector-based operational plan with more limited deployment activities. This is the case of the Evaluation of the National Plan for Transition to Digital Terrestrial Television (TDT) (AEVAL, 2009).

Figure 6. Main activities identified in the National Plan for Transition to DTT. Source: AEVAL (2009).

Group	Objective	Activities
DTT coverage activities	To implement DTT coverage percentages equal to the current analogue coverage in each Transition Project (TP) within the established deadlines	1. Available data on broadcasters and centres outside the agreement 2. Additional data 3. Technical analysis of the TP 4. Logistic support for the supply of transmission and installation equipment 5. Cost analysis 6. Maintenance 7. On-Off Planning 8. Coordination with the agents involved in this group of activities 9. Analysis of planning compliance 10. Post-analogue switch-off. Emergency management office 11. Drafting periodic monitoring reports
Activities aimed at users	To implement all the actions required for users to adapt their homes for DTT reception	1. Market research of the TP territory 2. Adapting collective antennae (Buildings) 3. DTT receiving equipment (Homes) 4. Stock Management of reception elements 5. Costs 6. Coordination with the agents involved in this group of activities 7. Analysis of the degree of DTT penetration 8. Post-analogue switch-off. Emergency management office 9. Drafting periodic monitoring reports
DTT information activities	Correct dissemination and user feedback management with regard to DTT and the National Plan for Transition	1. Drafting a specific Plan for communication and dissemination 2. Complaints and Feedback Management Office 3. Web Server 4. Monitoring dissemination activities 5. Costs 6. Coordination with the agents involved in this group of activities 7. Drafting periodic monitoring reports

Along with the identification of the deployment activities, the evaluator analyses their formalisation which, based on their scope and complexity, may constitute the document detailing the intervention, and which normally includes a series of measures that should make a coherent contribution to achieving the objectives proposed by the planners. In other cases, they are the documents on the formalisation of one or multiple activities, for example, specific action plans. Finally, the programming instruments may adopt the form of regulatory or legal mechanisms.

Example of regulatory or legal mechanisms. Evaluation of the Spanish Renewable Energies Plan or PER in Spanish (2005-2010) (AEVAL, 2011).

In the section that describes the intervention in the final evaluation report, it is mentioned that the "PER focused (...) on policy measures to regulate (the) (...) grants for the technological areas to flourish and at the same time be profitable for entrepreneurs, which would promote a productive network that would permit the development of different sources of energy: they have all developed unequally based on technology (...) with different levels of success attributable to very different causes. And obviously, the crisis (...) has affected the development of the measures".

The next step in the analysis of the deployment activities is the assessment of the coherence between the activities and the objectives, and between the activities and the outputs generated. Occasionally interventions are designed with objectives that do not contain measures or activities; or activities that generate outputs that have no (or very little) relation to the intended results. On other occasions, the evaluator may find that the measures are not aligned with any objectives or are defective with regard to their proportion.

Example of proportionality of organisational measures. Evaluation of the Plan for Measures to Improve Cross-Border Healthcare Services (AEVAL, 2017).

This Plan contains organisational measures, mainly on extending schedules and increasing and reclassifying personnel, where problems were detected in the results with regard to shortcomings in their coherence and proportionality. The primary measure consisted of achieving 24-hour availability of inspection services at certain borders, but it was found that the traffic of goods did not require this extension, rather what was required was greater flexibility in existing schedules. The evaluation concluded that "with regard to its technical assessment, the plan is an incoherent instrument which mixes a strategic focus in its formulation with an operational development that occasionally limits itself to partial measures that have little in common with the declared objectives and on other occasions it exceeds its scope, thus compromising the effectiveness of the public resources allocated to said measures".

Figure 7. Alignment between the objectives and measures of the Plan for Measures to Improve Cross-Border Healthcare Services. Source: AEVAL (2014).

PLAN FOR MEASURES TO IMPROVE CROSS-BORDER HEALTHCARE SERVICES		
Operational objectives	Measures	
A) TRANSIT OF GOODS: Border Inspection Posts (BIP)	- Organisational - Coordination - Buildings and facilities - Availability of technical and computer-aided means - New protocols	15 BIP of 42 in 2006 (first stage)
B) OTHER FREIGHT CONTROL SERVICES AND OTHER CUSTOMS ZONES	Not considered in the Plan	
C) PASSENGER TRANSIT: - International Vaccination Centres (IVC) - Emergency healthcare	- Organisational - Reinforced control - Mechanisms to ensure response capacity (management projects)	

In order to compile and validate the information required to analyse coherence, different techniques such as documentary research, interviews, surveys, or group techniques with the participation of key stakeholders are used.

Logical Framework Matrix

This coherence analysis may be represented with the logical framework matrix, applied to the implementation of the measures,

Figure 8. Diagram of logical framework matrix for the deployment of a measure belonging to an intervention. Source: Author's own.

MEASURE:				RESULT:			
	DEPLOYMENT INSTRUMENTS					AWAITED OUTPUTS	
	In-charge	Coordination	HR	Economic resources	Monitoring	Output	Indicators
Activity $_1$							
Activity $_2$							
Activity $_3$							
...							
Activity $_n$							

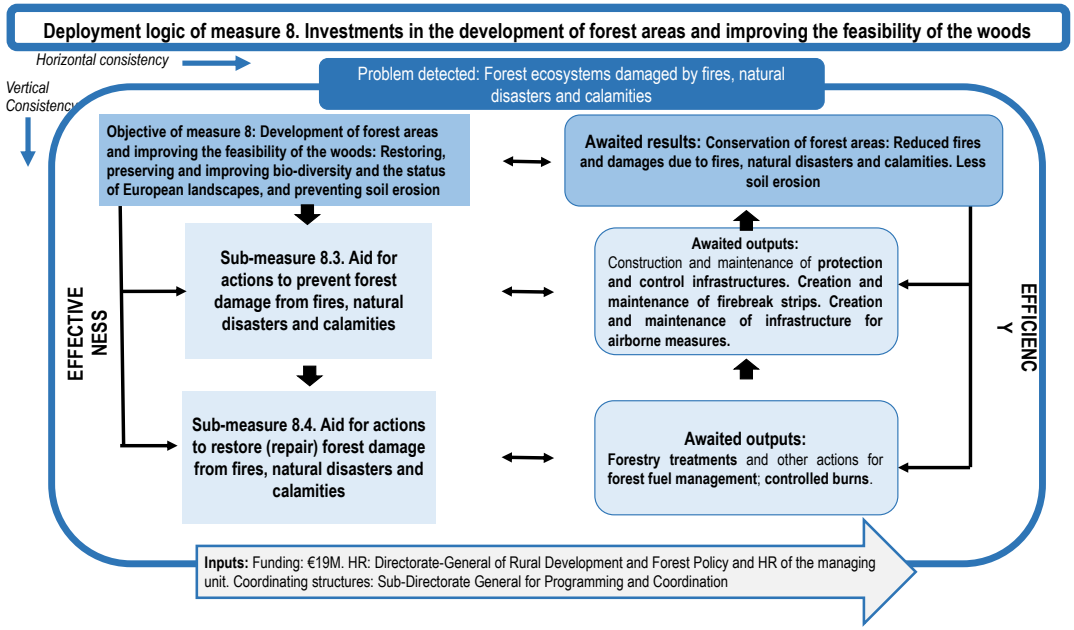
As may be seen in the figure, the upper part of the table indicates the measure (which in turn is linked coherently in its design[12] to a specific objective and a general objective with

[12] As indicated, this part of the comprehensive assessment is the goal of the Guide for Evaluating Public Policy Design (Institute for the Evaluation of Public Policies, 2020).

their awaited results and impacts respectively) and the expected result. The activities are the implementation of said measure, by applying different deployment instruments (those that are deemed necessary may be added) and the expected outputs of these activities, with their corresponding indicators of measurement (the table columns). These outputs will contribute to achieving the awaited objectives.

In the case of complex public policy interventions that contain actions where the measures are interventions in and of themselves, it is necessary to have a cascading logic of the intervention (logical sub-frameworks) of each deployment activity, as shown in the following figure included in the intermediate evaluation of the National Plan for Rural Development (AEVAL, 2017) where this tool was applied to each measure.

Figure 9. Deployment logic of a measure of the National Plan for Rural Development. Source: Author's own based on the AEVAL evaluation (2017).



In the Guide for Evaluating Public Policy Design (Institute for the Evaluation of Public Policies, 2020), we can find this same tool applied to the analysis of the theory of action.

Analysis of synergy and antagonisms

When the intervention is deployed by means of several activities/measures, they may be independent or may interact among themselves, either positively (synergy) or negatively (antagonisms), and thus influence the achievement of the objectives. That is to say, the measures may boost or obstruct according to their degree of influence or sensitivity. The evaluation must identify the most influential and sensitive ones that may have either a snowball or carry-over effect, in order to highlight the risks associated with defects in the execution of the deployment activities. This analysis is especially relevant in complex intervention evaluations, at the level of public policies or strategic plans, for the consequences that these effects may have on the success or failure among measures.

Synergy Matrix

An effective tool for this analysis is the ***synergy or interactions matrix***. By assessing the degree of interaction of each activity/ measure with regard to the rest, it analyses the independence, influence, and the sensitivity of each one as well as its level of synergy or antagonism. The scale of assessment is defined by the evaluator and either a qualitative or a quantitative scale may be used. In any case, it is necessary to always have the participation of the key stakeholders of the intervention, usually the persons or bodies in charge, or the managers, in order to obtain the assessment of the interactions.

In the matrix used to evaluate the National Programme for Rural Development 2014-2020 (AEVAL, 2017) a quantitative scale was used to assign different levels of effect between the different measures and to determine the level of influence on the achievement of specific objectives assigned to each one.

Figure 10. Scores assigned by the managers of the measures to assess the synergy between the sub-measures of the National Programme for Rural Development. Source: AEVAL (2017).

	SUB-MEASURES														
	M1.1.	M1.2	M4.2	M4.3.1.	M4.3.2	M7.8	M8.3	M8.4	M9.1	M15.2	M16.1	M16.1+16.2.	M16.2	M16.5	M16.6
M1.1.		3	3	0	0	0	0	0	0	0	0	0	0	3	3
M1.2	3		2	0	0	0	0	0	0	0	0	0	0	2	2
M4.2	2	2		0	0	0	0	0	0	0	0	0	0	3	3
M4.3.1.	0	0	0		0	0	0	0	0	0	0	0	0	0	0
M4.3.2	0	0	0	0		0	0	0	0	0	0	0	0	0	0
M7.8	0	0	0	0	0		0	0	0	0	0	0	0	0	0
M8.3	0	1	0	0	0	0		0	0	0	0	0	0	0	0
M8.4	0	1	0	0	0	0	0		0	0	0	0	0	0	0
M9.1	0	0	0	0	0	0	0	0		0	0	0	0	0	0
M15.2	0	0	0	0	0	0	0	0	0		0	0	0	0	0
M16.1	1	2	3	1	1	0	1	1	2	2		5	4	4	4
M16.1+16.2.	1	2	3	2	2	0	1	1	4	2	3		5	3	3
M16.2	0	0	0	0	0	0	0	0	0	0	0	0		0	0
M16.5	2	2	4	0	0	0	0	0	0	0	0	0	0		2
M16.6	2	2	4	0	0	0	0	0	0	0	0	0	0	2	

The results of this example table display a low level of influence, therefore, the evaluation concluded that the measures were considerably independent and there was insufficient complementarity between them for synergies to be produced by their interaction.

Planning and execution of measures

All public interventions must possess a sequential plan of the activities to be performed for their correct implementation, with the intermediate goals defined in order to ensure correct progress and execution within the stipulated period of time. This plan must be drafted before the implementation of the public intervention so it may be monitored, to detect intentional deviations or the non-implementation of corrective measures or modifications that are required for the intervention to achieve the objectives it seeks. Additionally, and as

a result of the complex and changing context in which the processes of formulating public policies take place, it is important to have a clear but flexible line of action that highlights the distribution of the activities and their deadlines.

A good activity planning contributes to the effectiveness of the intervention, to achieving the results and therefore the objectives.

In the implementation analysis, the evaluator will verify the existence of this planning, its compliance with deadlines and intermediate goals, the mechanism for monitoring the progress of the activities and the modifications to the timeline and its causes.

Timeline

A widely used tool to manage implementation is the timeline or the graphical representation of a set of activities under certain requirements and within a stipulated period of time. It is recommended that the timeline specify the persons or bodies in charge of monitoring the fulfilment of each activity, as this will make it easier to seek explanations if the measure is not executed within the awaited time and according to the expected percentage of completion

Figure 11. Example of activities timeline. Source: Author's own.

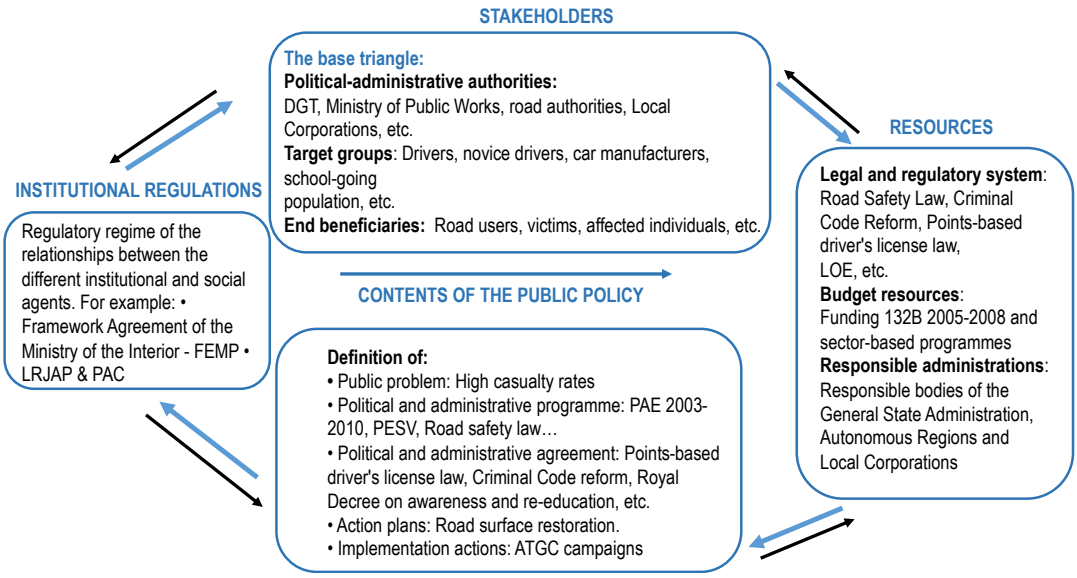
MEASURES	IN-CHARGE	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
Measure 1	Name, position									
Measure 2	Name, position									
Measure 3	Name, position									
Measure 4	Name, position									
Measure 5	Name, position									

1.2. Relevant stakeholders of the intervention

The "stakeholders" of an intervention are the individuals or collective groups that are in some way involved in said intervention and have different interests with regard to the objectives and results. Eric Monnier (cited in Ligero Lasa, 2016) establishes four categories of stakeholders: those responsible for the decision, those implementing the intervention, the social stakeholders that react to the intervention (beneficiaries, non-beneficiaries, supporters of a certain group) and the evaluating team itself.

Other authors such as Subirats, Knoepfel, Corine and Varone (2008) draw a distinction between administrative-political authorities, target groups, end beneficiaries, and other public or private stakeholders, individuals or organisations affected indirectly by the intervention (third-party groups). The following figure, used in the evaluation of the Strategic Plan for Road Safety 2005-2008 (AEVAL, 2009) represents this scheme alongside other essential elements of the implementation.

Figure 12. Elements of the analysis of the road safety policy, based on Subirats, Knoepfel, Corine and Varone. Evaluation of the Strategic Plan for Road Safety 2005-2008. AEVAL, 2009.



Regardless of the classification that is used, it is important to correctly identify the stakeholders and examine their positions with regard to the intervention. Each stakeholder group has different roles in the intervention, is differently related to each other, and their individual, occasionally conflicting, interests condition the implementation and subsequently the final results. The evaluation questions put to each group may be different based on the position defended by each with regard to the intervention.

Additionally, it must be considered that reality is dynamic and changing and therefore, stakeholders may modify their perceptions, they may have greater or lesser influence on certain aspects, and they may reinforce or change their positions. Therefore, it is important for implementation evaluation to have a certain degree of dynamism when identifying stakeholders in order to be aware of the changes that may gradually take place. Over the course of the evaluation process, new stakeholders may often appear as opposed to those that were initially identified, and they may be highly relevant for the evaluation.

The identification and analysis of stakeholders may be performed by different techniques and tools: focus groups, interviews, clustering based classification, map of stakeholders, network analysis, etc.

Stakeholder analysis is linked to evaluation criteria such as coherence, coordination, suitability, coverage, and participation, and provides information on evaluation questions related to the other criteria.

Stakeholder Matrix

Additionally, the position of the stakeholders with regard to the intervention helps the evaluator to understand the place of each stakeholder as well as their level of influence, expectations, information requirements, etc. To compile all this information, the evaluator may use a matrix that, owing to its content, is usually not included (at least fully) in the evaluations. Rather, it is a tool for the "internal" use of the evaluating team. The matrix lets us identify the different roles and manage the potential difficulties that may appear in the course of the evaluation tasks.

Figure 13. Stakeholders matrix of a public intervention. Source: Author's own.

STAKEHOLDER CLASSIFICATION	ROLE IN INTERVENTION	POSITION ON INTERVENTION	HIERARCHY / INFLUENCE ON INTERVENTION	ROLE IN EVALUATION	POSITION ON EVALUATION	NEED FOR INFORMATION
Responsible administration Other Administrations Private Institutions Social Agents Civil Society	In-charge Manager Beneficiaries End targets Users Interest group	For Against Indifferent/Neutral	High Medium Low	Commissioning Decision-making Collaborator Key informant	For Against Indifferent/Neutral	Defined for each stakeholder

1.3. Organisational instruments for deployment

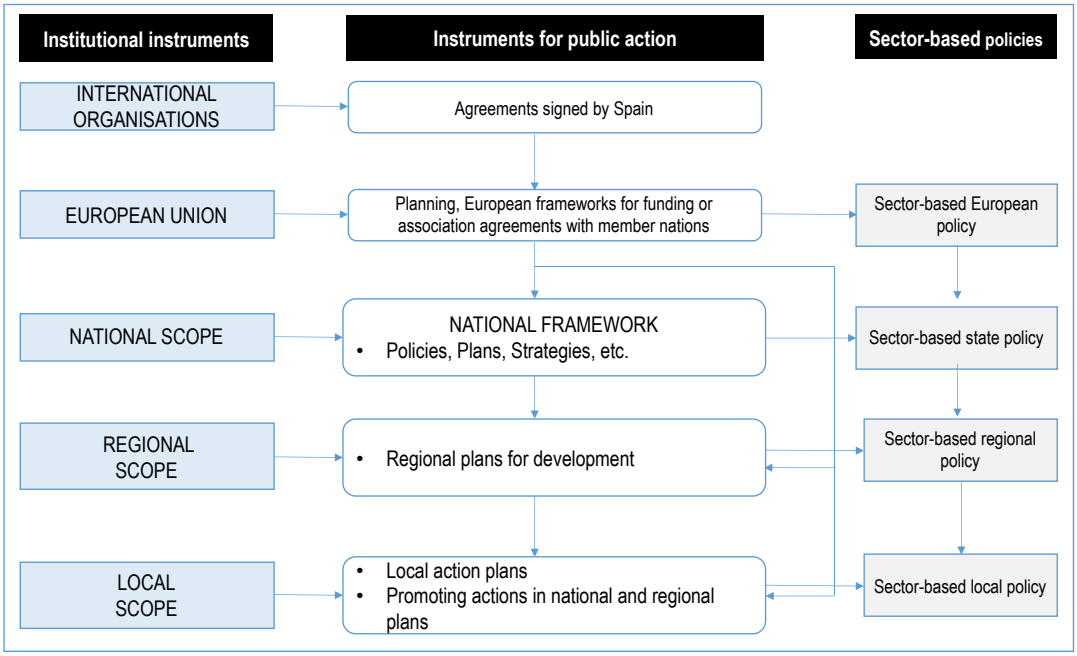
To evaluate the implementation of a public intervention, it is essential to correctly map the institutional and organisational structure and consequently, the decision-making system, the shouldering of responsibilities, how the execution of the different actions are coordinated, their monitoring, and the participation of the different stakeholders. This analysis helps us to acquire an in-depth knowledge of the intervention, entering fully into the study of the "black box".

Institutional organisation, distribution of powers and organisational structure

The framework of reference of the organisational instruments is the ***institutional organisation***. In other words, what levels of government participate in the implementation of an intervention, what is the concrete distribution of powers, and how are the management structures developed. Later we shall see the forms, if they exist, that are used to coordinate the different actions and ensure that they all come together to achieve the objective of the intervention, and to measure the progress made.

This analysis must consider whether public policies are currently becoming more complex owing, among other factors, to the greater cross-cutting nature of public actions, which increasingly approach more complex problems with greater interdependence. A second factor is the participation of a higher number of institutions and levels of government. Thus, when performing the implementation evaluation of a public policy in Spain, the existence of inter-related institutional instruments and the confluence of various sector-based policies must be taken into account, as is displayed in the following figure.

Figure 14. Diagram of institutional instruments and public action at different levels of government common in a decentralised nation such as Spain. Source: Author's own.



The relationship between the institutions and the executed policies is analysed to check the coordination and complementarity between them, as well as the degree of inter-dependence and their influence in achieving the objectives of the evaluated intervention. Although the evaluation uses the coordination criterion to assess all the relationships between the different institutional structures, it must be taken into account that, with regard to inter-administrative relationships, the law differentiates between coordination, collaboration, and cooperation[13].

It is especially necessary in implementation evaluation to analyse the ***distribution of powers*** between the levels of government, regulated in Spain by Articles 148 and 149 of the Constitution, the Statutes of Autonomy, and the organisational laws of local bodies, and which determine the actions that may be performed by some institutions or others, and the field of action of the stakeholders in a specific intervention.

[13] Article 140 of the Law 40/2015 of 1 October on the Legal Framework of the Public Sector, differentiates between the principles of coordination, collaboration, and cooperation. It refers to collaboration "as the duty to act with other Public Administrations in order to achieve common goals"; cooperation, "when two or more Public Administrations, undertake specific commitments for a common action"; and coordination, when "a Public Administration and uniquely, the General State Administration Services, has the duty to ensure the coherence of the actions of the different Public Administrations affected by the same topic for achieving a common result".

Figure 15. Distribution of State-Autonomous Region functions without transfers for coastal management. Source: Evaluation of the Management and Functioning of Coastal Demarcations to Protect Publicly Owned Marine Territory, from the perspective of its adjustment to the Water Framework Directive as well as the Marine Strategy Framework Directive. (AEVAL, 2012).

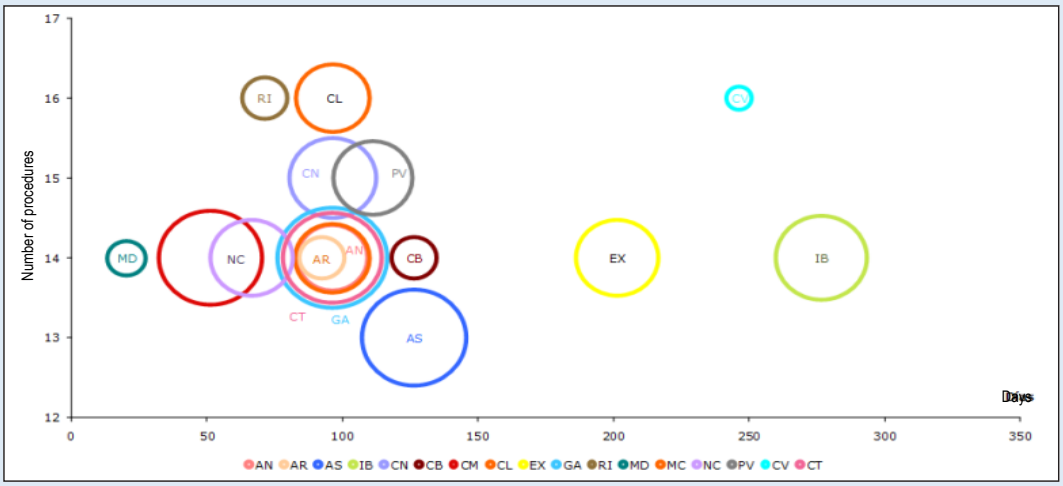
Distribution of functions (without transfer)	State	Autonomous Region
Within the Publicly Owned Marine Territory (DPMT)	Demarcation	To be heard
	Affiliation (recorded following the State's favourable report on the project and approval by the Autonomous Region)	
	Reservations	Mandatory report
	Recovery of illegal construction land	
	Concessions	Mandatory report
	Charges and fees	
	Penalty system	
	Public works	Mandatory report
Coastal protection zone (Zsp)	Mandatory report	Empowered (authorisations, charges and fees, surveillance and penalty system)
Land management and town planning	Mandatory report	Empowered

In the previous figure, in some cases a certain level of government may undertake an action (for example, the demarcations made by the General State Administration Services in Publicly Owned Marine Territory) but make another government body a participant (in the same example, listening to the affected Autonomous Regions) and in other cases, the powers are in the other direction (for example, in the case of land management and town planning, where the mandatory and binding reports are made by the General State Administration Services, but the power is that of the Autonomous Regions).

The distribution of powers may affect the implementation of the interventions by incurring different speeds that may lead to errors or inequalities in achieving the results in the affected territory, thus the importance of correctly identifying the distribution of powers that may explain the different intermediate and final results.

Example of how differences in powers at the local level and at the level of the autonomous regions, and their speeds of implementation may affect an intervention. Evaluation of administrative burdens in company creation (AEVAL, 2012).

The process of creation of an ICT company. Territorial comparison by Autonomous Regions. 2011.

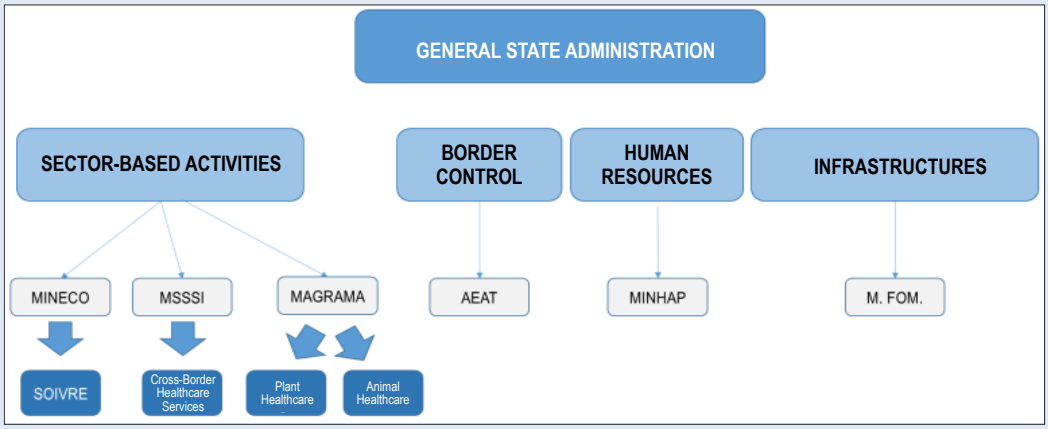


The evaluation states that "the situation in 2011 of the process of creation of ICT companies (...) in different Autonomous Regions shows that the majority is situated in the range of 14-16 procedures, with a certain dispersion with regard to the duration of the procedure, which may range between 50 and 150 days. If additionally, the cost of the procedure (reflected by the diameter of the circumference) is taken into account, the best global result is that obtained by the Region of Madrid, the one closest to the axis and with a small diameter. The regions that are farthest from the axis are Extremadura, Comunitat Valenciana and Illes Balears".

On the other hand, "it is evident that the administrative 'costs' are concentrated in municipal activities, in contrast to other state procedures or others whose regulatory development and ordinary management correspond to the Autonomous Regions".

Occasionally, even within the same level of Government, powers are distributed over different departments, which may be identified to consider the role of each one in the implementation of the evaluated intervention.

Example of the analysis of the distribution of powers in a sector-based policy at the same level of Government. Evaluation of the Plan for Measures to Improve Cross-Border Healthcare Services (AEVAL, 2014).



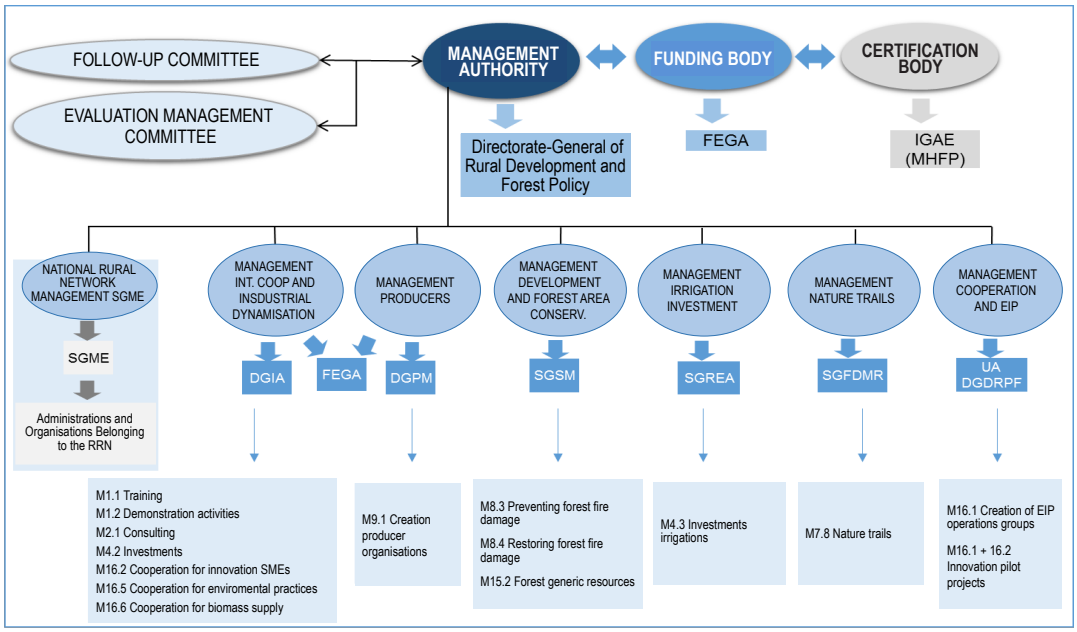
As may be observed in the figure, up to six different ministerial departments and organisations with powers of governance and the development of the policy under evaluation were detected, which increases the need for leadership. The evaluation report indicates that "in spite of the ample group of involved stakeholders, a lack of leadership is displayed" and that "each one of those involved perceives the problems and proposes solutions from their position and interests, and therefore, there are considerable contradictions in the Plan that continue to affect the correct functioning of the services" and it is recommended that "networking actions should be boosted as an element for effectiveness".

The ***structures*** in charge of fulfilling the intervention[14] are almost always derived from the distribution of powers. Firstly, the executive structures where the main decisions are taken must be described and analysed, both at the planning level[15] (objectives, intermediate goals, awaited results, monitoring and evaluation) and in the designation of the persons or bodies in charge of the implementation and development of the measures that seek to achieve the objectives, as well as the units managing said measures or activities through specific processes that possess certain resources.

[14] There are different implementation perspectives. The most well-known ones are the top-down and the bottom-up approach, but generally, they all refer to the need to mention the leaders, the persons or bodies in charge, stable structures, management and stakeholder support, and monitoring before the evaluation.

[15] This statement refers to the planning of the implementation of the intervention because, as we have mentioned, the stages of defining the problem and choosing the alternative have already taken place and are included in design evaluation, which is studied in another Guide (IEPP, 2020).

Figure 16. NPRD management structure. Source: Implementation Evaluation of the National Plan for Rural Development (AEVAL, 2017).



A Public Administration is an organisation based, among others, on the principle of hierarchy and therefore the chain of command must be analysed in an implementation evaluation that, as depicted in the previous figure, may cover several departments.

The shouldering of responsibilities is derived from this chain of command, although their planning and detailing may vary from one intervention to another. For example, the Evaluation of the Human Resources Quality Plan of the General Services of the Administration of the Autonomous Regions of the Balearic Islands (AEVAL, 2015), mentions that a working group was created for the implementation and monitoring of the Plan. This working group created records of each action, indicating the persons or bodies in charge (units or departments) of their execution, the objectives to be achieved on a yearly basis, and defining indicators to measure them.

For the implementation, it was necessary to design ***processes*** that systematically described the procedures to be followed by the units executing the measures. In the interventions executed by public administrations, these processes may be explicitly stated in the planning (for example, by means of the regulatory bases of a grant), be a part of the general functions of the unit (for example, mentioned in the royal decree laws of ministerial structures); or they may be implicit. In any case, one of the characteristics of implementation evaluation is the identification, and in some cases the analysis, of these processes where administrative units are especially relevant.

Example of process analysis. Evaluation of the Quality of Service in State Museums (AEVAL, 2008).

The evaluation ascertained that "communication is one of the key processes identified by all museums. However, this is not normally described or documented. Specific goals are not set, let alone indicators. Certain procedures, tasks, or activities, including the assignment of those responsible, are documented. However, there is no standard procedure amongst the museums evaluated".

Organisational Analysis

On certain occasions, it is necessary to undertake an organisational analysis, although it is generally used to evaluate the quality of services and not so much in public policy evaluation. There are different techniques and tools to perform these analyses, although the most well-known ones are the following:

The EFQM Model (European Foundation for Quality Management): The goal of this model for excellence is to help organisations to better comprehend themselves, to make an objective, rigorous, and structured analysis of their functioning and therefore, to improve their management. It analyses what it calls "enabling agents" (leadership, personnel, strategy, partnerships and resources, processes, products and services) and those linked to "results" (in personnel, in clients, in society and in key results) thus leading to learning, creativity and innovation, creating an system for continuous improvement.

The EVAM Model (Evaluation, Learning and Improvement): It has a simple and accessible methodology that displays the level of quality in organisational management and results, helps to perform an assisted self-evaluation which includes an initial analysis of the organisation's maturity and its level of provision of services; and shows the route to be taken, providing organisations with tools to improve their performance.

The CAF Model (Common Assessment Framework): It is a tool to manage the overall quality of an organisation, developed for and by the public sector and inspired by the Model of the European Foundation for Quality Management (EFQM). It is based on the notion that excellent performance results of an organisation, in citizens/clients, in persons and in society are achieved by means of a leadership that heads strategy and planning, personnel, partnerships, resources and processes. The model examines an organisation from different angles simultaneously, with a holistic approach to analysing the organisation's performance.

Mechanisms for coordination

As we have seen, the implementation of an intervention consists of highly complex processes where different institutional and administrative structures play a role and which require mechanisms for coordination in order to act in a homogeneous and coherent fashion, and increasingly, with the participation of the rest of the stakeholders that stand to benefit more

or less due to the intervention[16]. The mechanisms for coordination and participation may have different forms and scope, which require designing a network of actions that must come together (at least in theory) to achieve the objective.

In legal terms and with regard to the relationship between Public Administrations, Law 40/2015 of 1 October on Public Sector Regulation, defines the principle of coordination in its Article 140. According to this principle, "a Public Administration, and uniquely, the General State Administration, has the duty to ensure the coherence of the actions of different Public Administration affected by the same issue, in order to achieve a common result, where the Constitution and other legal systems so provide".

Applying this principle to the implementation of a public intervention, mechanisms for coordination are aimed at achieving the coherence and formation of the intervention to promote the achievement of the strategic goals as well as the creation of different types of synergies (of association and collaboration) to achieve specific goals.

These mechanisms may be quite varied, depending on the characteristics of the intervention, and they are evaluated according to criteria such as coordination, coherence, and complementarity:

- Partnership agreements between the General State Administration and the Autonomous Regions, for example, to define the participation and responsibilities of the parties in undertaking certain activities of the evaluated intervention. They may include funding and indicators for monitoring.

- The creation of specific units with the concrete goal of coordinating certain activities. For example, the centralised collection of information, decision-making, or dialogue with interest groups or stakeholders.

- Assigning coordination tasks to certain units with a special leadership role in implementing the intervention.

Sometimes, various departments or units involved in the coordination have simultaneous roles, as in the case of the evaluation of the National Plan for Rural Development, performed by AEVAL in 2017.

[16] Subirats, J., Knoepfel, P., Larrue, C. and Varone, F., Análisis y gestión de políticas públicas, Ariel, Barcelona (2008).

Example of instruments for coordination. Implementation Evaluation of the National Plan for Rural Development (NPRD) 2014-2020 (AEVAL, 2017).

It is explained in the evaluation that the "implementation of the NPRD involved the participation of different stakeholders (such as the managing authority, the managers of the measures, and the funding organisation). A Directorate-General was in charge of managing the NPRD through a Sub-Directorate that in turn managed a specific measure of the Plan and was also in charge of coordinating the authorities managing the different regional programmes for rural development. Mechanisms for coordination (both formal and informal) were established to ensure the quality and effectiveness of the programme, as well as different bodies in charge, although the fact that the management authority had multiple roles in the programme caused a degree in uncertainty in its partners, whether managers or the funding organisation.

Additionally, according to the report, coordination and communication with the Autonomous Regions within the NPRD varied, depending on the topics and the measures. Thus in some cases, there was no communication and in others, there was communication of varying degrees: from participation in different forums and committees where decisions were sometimes taken and sometimes not; to the unilateral communication of grants that prohibited double funding.

There may also be various consequences of a lack of coordination. Generally, they give rise to inefficiencies such as redundant petitions, different requirements for the same function, etc.

Mechanisms for participation

Public interventions increasingly possess different mechanisms for the participation of the most institutionalised stakeholders (managers, associations of possible beneficiaries or affected individuals, civilian bodies, etc.), especially at the planning stage.

In the case of drawing up laws and regulations, Article 133 of the Law 39/2015 of 1 October on the Common Administrative Procedure of Public Administrations, specifically regulates citizen participation with the proviso that "prior to the drafting of the project or bill or regulation, a public consultation must be performed through the website of the relevant Administration in order to compile the opinions of the citizens and the most representative organisations that are potentially affected by the future regulation[17], on:

a) The problems that the initiative seeks to solve.
b) The need and opportunity for its approval.
c) The goals of the regulation.
d) Alternative regulatory and non-regulatory solutions possible.

[17] The law also mentions some cases where these participatory procedures may be dispensed with.

In practice, and setting aside the legislative process, the mechanisms and bodies for participation are intended to provide counsel and advice, without being directly related to the decision-making process, restricted to the planning and design stage, without continuing on to the policy implementation stage.

Therefore, the drafting stage of interventions (consultation, public information, contributions and amendments and above all, similar to the provisions of Law 39/2015 for regulations) often entails a participatory process of variable intensity, but there are a fewer cases of participation in the implementation stage through formulas for cooperation or co-management, either as instrumental participation (delegated execution in associations, for example) or as co-execution and supervision of public action (Alvira, 2018).

This is in spite of the fact that experts agree that in general, and at any stage in the life cycle of public policies, "the two formal mechanisms that institutionalise citizen involvement in resolving public affairs are the regulatory framework and stable participation bodies" (Parés and Resende, 2009, cited in Pastor, 2013) and that with regard to public action, "citizen involvement in its design and implementation improves its effectiveness and legitimacy" (Ruiz, 2013).

In order to evaluate the mechanisms for stakeholder participation in the implementation of an intervention, the evaluator must take into account, in addition to participation, other criteria such as representativeness, inclusion, early involvement, flexibility, accessibility, independence, benefit or satisfaction, among others.

To be noted

It is important to distinguish between stakeholder participation in the implementation stage of a public intervention and stakeholder participation in the implementation evaluation.

*Participation in the **implementation stage of the intervention** refers to the degree to which the stakeholders are involved in the deployment and monitoring of the intervention.*

*Stakeholder participation in the **implementation evaluation** refers to the degree of involvement of the stakeholders and how they participate in the different stages of the evaluation process. Evaluations with a pluralistic approach, as adopted in this Guide, incorporate stakeholders into the evaluation process. The problems that may arise, such as an excess of specificities, debates focused on the short term (Subirats, 2001), or the exclusion of the most vulnerable groups (Ruano, 2010) occur at the time of managing stakeholder participation in the evaluation.*

Example of mechanisms for participation. Evaluation of the Strategic Plan for Road Safety 2005-2008. (AEVAL, 2009).

The evaluation report states that "the Directorate-General of Traffic was in charge of designing and drawing up the strategic plan, PESV 2005-2008 for which it was aided, not only by institutional agents who would later participate in the implementation and execution of the different programmes, but also by social agents. This participatory process was one of the hallmarks of the PESV".

Information to assess these criteria is obtained through documentary research, stakeholder interviews or any other social research technique, mainly qualitative, although a specific tool for quantitative assessment may be developed. Alvira (2018) considers public participation to be an inherently complex and value-laden concept, and points out that currently there are no agreed-upon evaluation methods or reliable instruments to measure it.

In any case, mechanisms for participation must be sufficiently stable, and have enough time, economic means and materials required to undertake the process (Department of Public Administration and Justice. the Basque Government, 2014), thus giving them the capacity to affect public interventions.

1.4. Resources allocated for the intervention

Adapting the resources to the policy or programme to be implemented is decisive for the latter's success, especially when dealing with economic resources, contributing to its improved effectiveness and efficiency. On one hand, if the amount of resources required has been underestimated, then all the planned measures cannot be implemented, and on the other, if they have been overestimated, it is a lost opportunity, since these resources could have been used for another policy, an important issue in contexts of scarcity.

The evaluator must assess whether the resources (economic, personal, material, and technological) allocated for the intervention are sufficient and are available in the correct form and sequence to obtain the outputs with which to achieve the proposed objectives. To perform these analyses, it will be necessary to obtain information by researching key documents, monitoring and execution reports, interviews of persons or bodies in charge, managers, and when applicable, stakeholder surveys.

Economic resources

In an ex-post evaluation, resource analysis is useful in accounting for the results obtained, especially when they are not the awaited ones.

Analysing economic resources is important in implementation evaluation in order to identify the risks associated with the pace of development of the actions and to take the measures required to correct possible deviations before the results of the intervention are committed. Specifically, special attention must be paid, among other questions, to:

- ***The degree of sufficiency*** of the economic resources to commit to the planned resources, with regard to the amounts, nature, as well as availability. The deficiency of economic resources may lead to unperformed activities and depending on their level of criticality, the result of the intervention may be compromised in terms of effectiveness and efficiency. The sufficiency analysis must also pay attention to the evolution of the assigned economic resources, beginning from the planning and throughout the implementation, in order to identify possible deviations that may compromise its correct execution.

- ***The degree of suitability*** of the economic resources. Their selection may generate inconsistencies with the intervention design, thus affecting the level of coverage of the target population. For example, in an aid programme, it is not only necessary that the allocated budget be sufficient but the mode of providing aid (grants, subsidies, credit, direct investment, transfers...) that is the most suited to the target population in order to achieve the set results and objectives, should also be taken into consideration.

- ***The distribution and weight*** of the economic resources lets us assess the internal coherence of the measures, that is to say, whether their planning is in balance with the critical nature of the activities/measures for achieving the objectives.

Example of resource analysis. Intermediate Evaluation of the National Plan for Rural Development 2014-2020 (NPRD) (AEVAL, 2017).

This Plan had a budget of 435 million euros and was aimed at boosting rural development in Spain by financing activities (mainly aid programmes), grouped into eight measures and 15 sub-measures. The evaluation showed that 58.69% of the budget had been destined to a single measure (M4.2 in the following table) and therefore, the difficulties in its implementation were a crucial impediment for the NPRD to reach its target completion objectives in the intermediate years.

MEASURES	Total public contribution	EAFTD Contribution	Weight (%Total public spending)
M1.1. Aid for professional training and skill acquisition	1,862,655.00	1,490,124.00	0.43
M1.2. Aid for demonstration activities and information actions	1,862,655.00	1,490,124.00	0.43
M4.2. Aid for investment in transformation/marketing and/or development of agricultural products	251,781,934.00	133,444,425.00	58.69
M4.3. Aid for investments in irrigation infrastructures	31,864,291.00	16,888,075.00	7.43
M7.8. Creation and maintenance of Nature Trails whose route runs through two or more Autonomous Communities, or in island territories	18,743,703.77	9,934,163.00	4.37
M8.3. Aid for actions prevent forest damage from fires, natural disasters and calamities	19,415,552.83	10,290,243.00	4.53
M8.4. Aid for actions to restore (repair) forest damage from fires, natural disasters and calamities	3,000,000.00	1,590,000.00	0.70
M9.1. Aid for creating groups and organisations of producers in the agricultural and forestry sectors	13,038,587.50	10,430,870.00	3.04
M15.2. Aid for the conservation and promotion of forest genetic resources	4,028,000.00	3,021,000.00	0.94
M16.1.Aid for the creation and functioning of EIP operational groups for agricultural productivity and sustainability	4,685,925.11	2,483,540.29	1.09
M16.1+2. Aid for EIP operational group projects for agricultural productivity and sustainability	42,173,325.90	22,351,862.71	9.83
M16.2 Aid for pilot projects and the development of new products, practices, processes and technologies (food industry)	16,143,011.00	12,914,408.80	3.76
M16.5. Aid for environmental projects and practices for associative integration	1,241,770.13	993,416.10	
M16.6. Aid for the sustainable supply of biomass for associative integration	1,241,770.13	993,416.10	0.29
M20.1. Technical Assistance	2,855,005.66	1,513,153.00	0.67
M20.2. National Rural Network	15,094,339.62	8,000,000.00	3.52
TOTAL	**429,032,526.65**	**237,828,821.00**	**100.00**

- The ***overall distribution*** and ***weight of the economic resources*** with regard to the resources allocated to other interventions with similar objectives or with regard to other powers that may be attributed to the body in charge, which lets us assess questions of complementarity of the measures with those of other interventions.

- The ***budget or investment effort***, in terms of the contribution made by the measures to achieving the objective of the intervention, that may be used as an indicator to analyse the results.

- The ***real destination of the economic resources***, that is to say, whether the resources are being used for the target for which they were planned and allocated.

- The ***pace of execution*** of the budget. The implementation evaluation must identify and assess the risk associated with the budget execution according to the projected timeline and the consequences of a deviation on the results of the intervention. This point is explained in greater detail in the following section.

Human resources

The personnel available for the intervention are essential for its success, both with regard to quantity and their ability or suitability for the required tasks. Reviewing the persons in the organisation who are charged with the implementation of the measures from a management point of view is highly recommended at all stages of the intervention and in all types of evaluation, but especially in implementation evaluation.

The analysis seeks to prevent unwanted effects in the black box of the intervention arising from problems in the human resource composition and characteristics, such as bottlenecks, delays in reaching deadlines, or the impossibility of performing the activities in the scheduled time and manner. In any case, it is worth remembering the complexity of the functions performed by the units in order to assess them in relation to the development of the commissioned activities.

They must therefore analyse questions such as the following, tailored to each intervention:

- ***The existence of sufficient and suitable human resources***. A low allocation of personnel for the implementation of an intervention may lead to delays, reduced quality in the provision of services or goods, or the need to subcontract, outsource or commission its management to other bodies. Additionally, if the personnel in charge of implementing a specific measure are not skilled or do not have the necessary experience to perform the tasks, it may also lead to problems of quality, delays in completing deadlines or ensuring legal security. These weaknesses almost always require greater efforts for training, in coordination with the distribution of concrete and specific instructions for each case.

- ***Establishing the dimensions of the tasks to be performed***. It must be assessed whether the task to be performed by the personnel is complex but limited in number, or relatively simple but numerous (processing many records, for example), or of a high frequency. Seasonality is also an important factor in some cases, as an increased workload may endanger the achievement of the objectives.

- ***Resources available to the personnel***. Occasionally, some technical or software tools are decisive in simplifying the workload of the persons that perform management or implementation tasks. This entails providing the necessary training for their use and the availability of economic resources to create or access them.

Example of human resource analysis. Evaluation of the management and functioning of the Hydrographic Confederations (AEVAL, 2009).

One of the objectives of the evaluation was "to check the degree of functional, institutional and resource-related readiness of river basin organisations to perform the new set of tasks that were commissioned by the state regulations established by the Water Framework Directive (WFD)". For this "a detailed analysis was performed of the organisational structure, the powers and functions performed by their organisational units, and the management and functioning of both river basin organisations, with special focus on certain administrative procedures". After the analysis, it reached the following conclusions and recommendations:

- Human resources had not experienced a process of adaptation commensurate with their increased functions, which led to extremely frequent outsourcing, which had been criticised by the General Comptroller of the State Administration (IGAE).

- The List of Work Positions (RPT in Spanish) was revealed to be insufficient in the evaluated confederations, both from the quantitative point of view and in terms of their adaptation to the new characteristics of the positions arising from the WFD's requirements.

- It recommended the modification of the personnel selection procedure and to incorporate personnel with other profiles (degrees and experience in biology or ecology...) that were different from current ones, with a majority of the bodies related to the management of water infrastructures. It showed that the efforts made by the evaluated confederations in this regard had not been successful, and when offers for jobs with these new profiles were published, they remained vacant.
- Finally, it highlighted the need for specific training on the WFD and its consequences for the functioning of the hydrographic confederations.

Technological and material resources

The implementation of the measures of an intervention is increasingly dependent on the available technological resources, both with regard to the internal management of the units, and the monitoring of the activities performed with regard to the target group that benefits from these measures. For this it is necessary to take into account digital applications, secure and inter-operational databases[18], accessible online[19] surveys, and even physical resources for internal management such as computers, suitable servers, access to management programs and applications, etc.

We must also remember that many of the interventions arising from programmes funded or co-funded by the European Union are directly planned and developed in online applications designed by the community management authorities.

In evaluations where the intervention requires specific material means, their suitability and the consequences of their deficiency for the results shall be studied. This is the case for means of transportation to perform deployment activities of the intervention, for example.

1.5. Instruments for monitoring, evaluation, and accountability

One of the most common flaws of public interventions is that they lack a system for data collection in order to ensure correct monitoring and evaluation. Information systems are systematically compiled datasets (AEVAL, 2015) that feed the monitoring system. This data comes from primary (directly compiled by the evaluator or is unprocessed) or secondary (data from other sources or processed) sources. Another frequent possibility is that there is a system for information and monitoring, but it is either not well-scaled or contains projections and indicators that are so open and general that they do not provide sufficient information to undertake the monitoring and evaluation.

[18] In some cases, it will be necessary to comply with the regulations on citizen access to information networks, digital management, survey formats, when applicable, and Law 19/2013 of 9 December, on transparency, access to public information and good governance.
[19] For example, in a subsidy which is applied for and managed by means of a software application.

In order to fulfil their objective, ***monitoring systems*** must be based on a prior quantification of objectives (also called standards) and indicators, as "otherwise, the degree to which the original objectives are being fulfilled cannot be evaluated" (European Commission, n.d.). Their main goal is "to explore and permanently analyse the degree to which the performed activities and the results obtained fulfil the planning goals, for a timely detection of eventual differences, obstacles or requirements for adjustment in planning and execution (SIEMPRO, UNESCO, 1999). This system must have the following characteristics (AEVAL, 2015: 81):

- Information compilation must be systematic and periodical (defined in planning), a person or body in charge of data collection must be appointed, with common directives that are known to all, and have a homogeneous and interoperable structure for data collection.

- Normally, the information is collected as indicators. Indicators are data or sets of data that help to objectively measure the evolution of a process or an activity. They must be useful, reliable, accessible (with regard to cost as well as effort), easy to interpret, comparable, relevant, timely, objective, precise, pertinent, sensitive, and constant over time.

- There must be an indicator baseline. It refers to the value of each indicator during a certain time which is established as the starting point to define the targets to be achieved and to assess their monitoring during the implementation phase. It is deemed to be the beginning of the intervention and the indicators are the reference or point zero.

- The sources of information must be specified.

All of these questions must be taken into consideration from the planning stage onwards, but it is essential in the deployment that there are no doubts regarding how, when, whom, and to whom should this data be reported. The best way to ensure this would be to include a section along with the monitoring and evaluation plan in the programming document, with detailed information on the system of indicators, information and monitoring, as well as the deadlines for submitting monitoring reports, the resources to be used, and the person or body in charge of drafting them.

Example of monitoring plan. Evaluation of the implementation of the National Programme for Rural Development 2014-2020 (NPRD). (AEVAL, 2017).

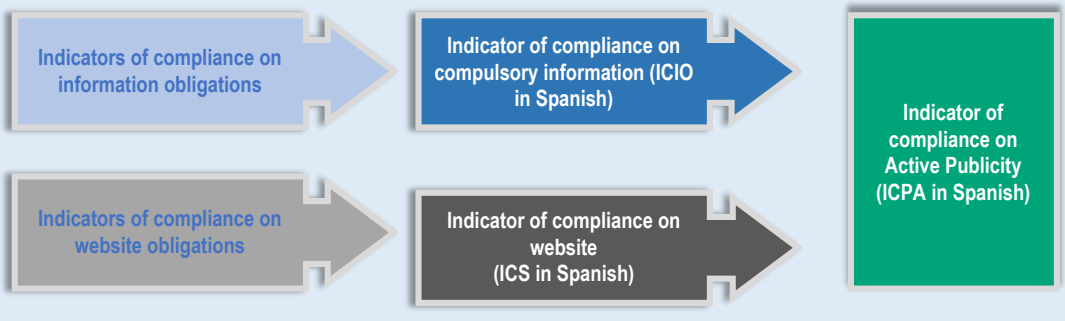
The evaluation indicates that the NPRD contains a series of indications for the monitoring and evaluation of the Programme. Specifically, "to provide 'suitable and timely information' to those responsible for decision-making, and related to the information requirements of the management authorities, the European Commission, and other parties interested in the evaluation". At the same time, the monitoring and evaluation plan "seeks to ensure the availability of the necessary data for the annual reports demanded by the European Union and the ex-post evaluation of the NPRD". It also identifies the persons or bodies responsible, and the organisations participating in the monitoring and evaluation system.

This notwithstanding, in practice the evaluation team faced several difficulties in this regard, of which a significant aspect highlighted in the evaluation report was "the low consistency and quality of the database". Occasionally, it is even "unknown how the datum included in the record provided by the managing authority was obtained". Besides, the additional programme indicators lacked a "methodology for calculation" and the "existence of records of these indicators" was not accredited, nor "how the set objective was calculated".

When an information and monitoring system has not been considered, or if it is not suitable, an ad hoc system must be built at the time of evaluation. It is a complex task that may present difficulties, above all, related to the non-existence of data, either because its collection was not scheduled in planning, or due to a lack of their availability at the time of evaluation. In these cases, the evaluating team must attempt to approximate as closely as possible the reality of the implementation from the data available and the limitations imposed by the lack of data to respond to the evaluation questions.

Example of building an indicator system. Methodology for the evaluation and monitoring of Transparency in Public Activity (MESTA) (AEVAL, 2016).

In order to monitor transparency, a system of indicators was designed, as Law 19/2013 of 9 December, on transparency, access to public information and good governance, did not mention indicators, but general principles and specific contents. The figure below shows the process of creating one of the indicators used.



The report explains that "the ICPA is obtained by adding the indicator of completion of the compulsory information (ICIO) and the supporting indicator of completion (ICS). The mathematical expression is the following:

$$ICPA = ICIO \times p1 + ICS \times p2$$

p1 and p2 being the weights of each component of active publicity, and fulfilling p1 + p2 =1".

The absence of information, indicators, and monitoring systems is not an issue to be taken lightly. As a matter of fact, "one of the main concerns of the European Commission regarding the evaluation of programmes that it funds (Structural Funds, ERDF, Cohesion Funds...) is precisely the establishment of a series of common standards adopted by all the States when assessing these programmes[20]" (AEVAL, 2015).

The intervention must also consider the existence of an ***evaluation plan*** that includes the evaluation activities to be performed, the resources, and the persons or bodies in charge.

Example of monitoring and evaluation systems. Evaluation of the Plan for Measures to Improve Cross-Border Healthcare Services. (AEVAL, 2015).

The report mentions that "in point 5, the Plan indicates the performance of an evaluation, 'once Plan 2006 has been executed', by AEVAL and commissioned by the responsible Ministries" and whose goal is to assess "the quality of the services provided by the General State Administration in the pursuit of its duties with regard to cross-border healthcare services", as mentioned in the evaluation report.

The monitoring and evaluation activity is usually included in reports that are systematically validated at different periods. Usually, the intervention establishes a ***monitoring and evaluation***[21] committee that is in charge of, supervises and validates, as applicable, the monitoring and evaluation reports. These monitoring committees may be in different locations (organisations, bodies, etc.) and have different compositions and levels of responsibility.

Example of mechanisms for monitoring. Evaluation of the Holiday Programme for Senior Citizens and Social Thermalism. AEVAL, 2010.

This evaluation states that "the monitoring committees –consisting of labour unions, the awarded company representing the business sector, the regional governments, the labour administration and IMSERSO– receive the reports drafted from a standard model which is distributed by the awarded company at every hotel participating in the programme".

[20] Since 1992, the European Commission has been working to develop a common system of indicators.
[21] The designations of these groups may vary. The same group may also be responsible for performing both activities.

In an age when citizens increasingly demand more information on public actions, implementation evaluation assesses the ***mechanisms for accountability included*** in the planning. These mechanisms may range from the publication of the planning process, monitoring and evaluation reports, to organising activities to disseminate information on the targets achieved or the commitment to the stakeholders (generally, the citizens).

1.6. Instruments for communication and dissemination

The instruments for communication and dissemination provide knowledge on the intervention and its measures, in two dimensions. The internal dimension deals with the management units or those affected by the implemented activities or measures, allowing greater fluidity in their shared channels of communication. It may include training activities, instructions with regard to a specific action, the creation of a common database, participating in periodic meetings, etc. It involves using tools for coordination to optimise the implementation.

Example of information and publicity strategy of an intervention. Internal and external dimensions. Implementation Evaluation of the National Plan for Rural Development (NPRD) 2014-2020 (AEVAL, 2017).

The evaluation report mentions "the NPRD information and publicity strategy was presented to the Monitoring Committee and the European Commission in November 2015 and (...) the opinion of the different units managing the measures was taken into account in its drafting. The information and communication activities range from exchanging internal information to publicity, launching a website or participating in different events or creating procedural guides that are defined in annual plans that also serve to monitor and evaluate the activities undertaken the previous year. These actions are meant for different target audiences: the general public, the potential and real beneficiaries, the management units themselves and other public administration targets". An Action Plan 2016 was also drafted with the information and publicity activities to be undertaken throughout this year. Finally, "instructions related to information and publicity aspects were released, for the managers of these measures".

In spite of these provisions, the evaluation declared that "the information and publicity strategy is lacking. While (...) there is some development with regard to the NPRD in general, this is not the case at the level of the measures, or it is incomplete", thus affecting the implementation of the measures.

The other dimension is external and it allows the stakeholders, especially the targets of the intervention, to learn about the existence of these measures, to participate and benefit from them, avoiding biases in the target population. Additionally, it also serves as a tool for transparency and suitability of the intervention.

Example of evaluation of the tools for communication and publicity to the targets of the intervention. Evaluation of the Programme to Aid Reindustrialisation Actions (REINDUS) (AEVAL, 2011).

The evaluation concluded that "achieving the REINDUS objectives depended on the participation of the (public and private) bodies in the programme. It is obvious that the degree of awareness of the target population is crucial for their participation, therefore the established mechanisms for dissemination in this regard are highly important, as an irregular dissemination or one that is not directed selectively to the sectors of greater interest for the programme objectives may introduce biases. Therein lies the importance of a correct dissemination. The REINDUS mechanisms for dissemination consist of isolated publicity sessions in territories to spread information on the programme, in coordination with the managers of other aid programmes of the Ministry of Industry, Tourism and Trade, or in collaboration with the Chambers of Commerce or associations. These actions are not widespread nationally, their coverage depends on specific requests made by certain regions and the allocation of resources by the Directorate-General of Industry. With the exception of 2010, there was no evidence of planning with regard to the sessions in the evaluated period". Recommendation: "It would be appropriate to improve the mechanisms for the external communication of the programme by implementing a dissemination plan that employs different channels of information, boosting regional and trade channels (Chambers of Commerce and business associations), and ensuring an overall reach, thus improving the egalitarian nature and transparency of REINDUS".

Both (internal and external) dimensions with their different activities are grouped under the communications strategy that must consider "the different targets and their information requirements, and from there, establish the most suitable channels of communication in each case (traditional as well as based on new technologies) for the greatest coverage" (AEVAL, 2015). It must include policy makers, managers, key stakeholders, and other interest groups and society in general, in an attempt at transparency and accountability.

Example of evaluation of the tools for communication and dissemination of measures. Evaluation of Administrative Burdens in Company Creation (AEVAL, 2011).

"The Royal Decree Law 13/2010 (...) is an advancement in offering entrepreneurs company structures that are easier to create and have economic and filing advantages. There are instrumental causes behind the still-developing stage of this route, such as the inability to execute any procedure by any other route if this one is selected, or the lack of an information and dissemination campaign on its advantages targeting entrepreneurs", states the evaluation report. Based on this conclusion, the evaluation recommends that "it would be appropriate to execute a publicity campaign on Royal Decree-Law 13/2010 to raise awareness among entrepreneurs on this company creation offer, its conditions and advantages, so they may select the legal route that best suits their needs".

2. ANALYSIS OF THE IMPLEMENTATION

Analysing the level of implementation, in an intermediate evaluation, seeks to assess whether the intervention is progressing correctly towards achieving the end results, identifying the deviations, divergences and delays with regard to planning, as well as their causes, which may be internal or external to the intervention. It also allows us to assess the projected deadline for reaching said objectives, as well as to identify risks derived from the design or implementation and to establish measures to mitigate them, if necessary.
In an *ex-post evaluation*, the analysis of the implementation serves the same purpose, although usually, its greatest potential lies in identifying the causal relationships between the process itself and the final results obtained.

Example of implementation analysis. Evaluation of the Plan for Measures to Improve Cross-Border Healthcare Service. AEVAL, 2014.

With regard to the implementation of the most relevant measure of the evaluated Plan, which consisted of extending the schedules for certain border inspection posts (BIP), the evaluation report explains that based on a framework agreement applied at the territorial level, the implementation of the measure led to "failures in the provision of the services". "There are evident failings in the distribution of BIP personnel's work schedules and in the system of hourly compensations. Both have a negative effect on the service provided at peak hours of goods traffic (...). From the personnel's point of view, this failing favours some personnel to the detriment of others, which arises from the labour advantages attached to the time extension, and which affects the quality of the service provided with delays in the inspection process on certain occasions". It concludes, therefore, that "the measure (...) may be the reason behind the increase in the time required for the passage of goods, owing to the personnel compensation systems established on the basis of the framework agreement", which is an undesired effect.

An added utility of this analysis is that it allows us to identify the errors caused by implementation defects which may be judged by criteria normally used in design evaluation, for example, coverage. It may be that an intervention is unable to achieve its results owing to a failure to reach the target population and the cause lies, not in the intervention design but in defective implementation, processes, or in errors in the mechanisms for communication or dissemination, thus obstructing a correct execution of the deployment activities or measures.

Example of errors in implementation that may affect coverage. Evaluation of the Programme to Aid Reindustrialisation Actions (REINDUS) (AEVAL, 2011).

This evaluation demonstrated that information regarding this aid programme did not reach all the collectives that were eligible to apply for it, owing to a defect in the programme dissemination strategy, thus affecting its level of coverage. The report indicates that "information regarding the programme is distributed through two channels: the publication on grants made by the public information systems of the Ministry of Industry, and through targeted dissemination in the territories meant to receive the aid". While the first functioned properly, the second demonstrated that "the publicity had greater effect in certain territories and during specific application periods. This latter aspect may be improved upon. Especially if we consider that there appears to be a certain relationship between the channel of information and obtaining the REINDUS grant. Territories where dissemination activities were performed submitted more applications, although this may also be attributed to the fact that said actions were performed during the regional deadlines for receiving applications". It recommended that "it would be appropriate to improve the mechanisms for the external communication of the programme by implementing a dissemination plan that employs different channels of information, boosting regional and trade channels (Chambers of Commerce and business associations), and ensuring an overall reach". This would improve, concludes the evaluation, the coverage, the egalitarian nature, and the transparency of the programme.

The degree of completion of the different stages of deployment of the initiative is usually calculated in terms of execution deadlines or intermediate results that constitute the intermediate goals.

Checklist

Execution deadlines refer to the degree to which the activity/ measure has been implemented and the intermediate results refer to the changes in behaviour, status, attitude or attestation by the beneficiaries as the planned activities are performed.

For this analysis, it is recommended to draft a checklist of measures that reflects the level of completion attained. This tool will allow us to perceive, with regard to the planned time and quality, if there has been a deviation in the implementation with regard to the plan, and the registered level of deviation.

Figure 17. Example of checklist for the completion of goals of a series of activities. Source: Author's own.

Measure	Activity	% of execution anticipated in stage t+1	% executed in stage t+1	Deviation from expectation
Measure 1	Activity 1.1	50%	20%	**-30%**
	Activity 1.2	80%	75%	**-5%**
	Activity 1.3	75%	75%	**0%**
Measure 2	Activity 2.1	60%	20%	**-40%**
	Activity 2.2	12%	23%	**11%**
Measure 3	Activity 3.1	94%	80%	**-14%**

Implementation evaluation must include an ***execution projection*** of each measure along with its corresponding indicators that allow us to assess the efficiency of the programme in reaching the intermediate goals (or some of them), and in completing the objectives within the scheduled date in the life cycle of the intervention. This exercise lets us identify critical points that require decisions to be taken on the design and implementation of the plan, in order to overcome the detected difficulties, when applicable.

In order to take these decisions, ***the criticality of each measure*** must be considered. It assesses the relevance of each indicator with regard to completing the objectives. Essentially, it is a methodology that allows us to hierarchically arrange measures and objectives based on their global impact, which helps us to take the right decisions and lets us direct our efforts and resources to the areas where they are most required and/or which must be improved.

Along with criticality, an analysis of the managers' ***confidence level*** with regard to achieving the projected results or the intermediate goals of the initial plan is also recommended. The analysis may be performed by assigning each measure a score between 1 to 5. The result obtained will depend, not only on the trajectory of the programme implementation until that moment, but also on the more or less conservative attitude of the managers consulted.

The data may be collected in a table that lets us view the measures, the associated indicators, the objective they must contribute to, their awaited value at the end of the intervention, the degree of execution at the moment of the intermediate evaluation; the execution projection (at different periods), the confidence of the stakeholders with regard to the projection and the relevance of each indicator in fulfilling the objective (in a weighting that may range from 1 to 5, for example), as may be seen below.

Figure 18. Projection of execution and completion of objectives according to measure, confidence level and criticality. Source: Author's own from the Evaluation of the Implementation of the National Programme for Rural Development 2014-2020 (AEVAL, 2017).

MEASURE	INDICATOR	OBJECTIVE	PROJECTED VALUE 2023	EXECUTION DEGREE 2016	EXECUTION PROJECTION cumulative % with regard to			CONFIDENCE IN PROJECTION (1-5)	RELEVANCE FOR OBJECTIVE (1-5)
					IN 2018	IN 2020	IN 2023		
M1.1	Total public spending on measure M1.1.	3A	1,862,655	0%	13	65	100	4	2
	No. of modules related to energy efficiency.	5B	2	0%	0	50	100	3	2
	No. of training days held.	3A	5,000	0%	10	40	100	3	3
	No. of participants in training activities.	3A	400	0%	20	60	100	3	3
M1.2	Total public spending on measure M1.2.	3A	1,862,655	9%	30	74	100	4	2
	No. of activities/operations subsidised	3A	20	65%	50	75	100	3	3
	Percentage of subsidised activities related to energy efficiency	5B	0.20	0%	50	100	100	3	2
M4.2	Total public spending on measure M4.2. in €	3A	251,781,934	1%	13	50	100	4	5
	Total investment (public and private) in €	3A	621,454,835	1%	13	50	100	4	5
	No. of operations subsidised.	3A	120.00	3%	40	80	100	3	3
	No. of Priority Cooperative Associations subsidised.	3A	20.00	15%	30	60	100	3	3
M4.3	Total public spending on measure M4.3. **(under 2A).**	2A	15,932,146	0%	30	50	100	3	5
	Surface **equipped** with irrigation infrastructures **(2A)**	2A	1,328	0%	30	50	100	3	5
	Total public spending on measure M4.3. **(under 5A).**	5A	15,932,145	0%	5	30	100	3	5
	No. of operations subsidised in M4.3.**(5A)**.	5A	3	0%	33	67	167	3	2
	Total surface area in hectares M4.3 infrastructures **(5A)**	5A	1,992	0%	5	30	100	3	5
	Percentage of irrigated land that has shifted to a more efficient irrigation system (modernisation)	5A	0.07	0%	5	28	100	3	5

In some interventions, there may be ***consequences of a failure to fulfil*** the intermediate goals or obligations. It is a common practice in plans and programmes co-funded by the European Union that are given a degree of flexibility when commencing the execution, however if they fail to correctly fulfil the tasks assigned to a certain goal deadline (rule N+3), they lose a percentage of the community funding. This may lead to problems in the implementation and final achievement of objectives, as contemplated in the aforementioned Evaluation of the Implementation of the National Programme for Rural Development 2014-2020 (AEVAL, 2017).

Example of problems in the deployment instruments and the consequences of failing to reach target deadlines. Evaluation of the implementation of the National Programme for Rural Development 2014-2020 (NPRD), AEVAL, 2017.

"The accumulation of EAFRD funds that must be executed owing to rule N+3 and the impossibility of the measure for executing them would make it necessary to modify the NPRD to distribute the excess among other measures that are able to absorb and execute them. (...) To this general problem affecting the entire NPRD may be added implementation problems owing to the design of the measure itself. The connection of the measure with the acknowledgement of new Priority Cooperative Associations (EAP in Spanish) and the difficulties faced by cooperatives to comply with the requirements (the minimum billing volume required by the sector and the cross-regional character are the major obstacles) prevent a greater level of execution. If the errors in the deployment instruments of the organisational integration policy of the Ministry of Agriculture, Fisheries, Food, and Environment (MAPAMA) are not rectified in time, they may compromise the level of execution of Measure M4.2 and endanger the fulfilment of all NPRD objectives, owing to the relevance of the measure in the

Risk Analysis

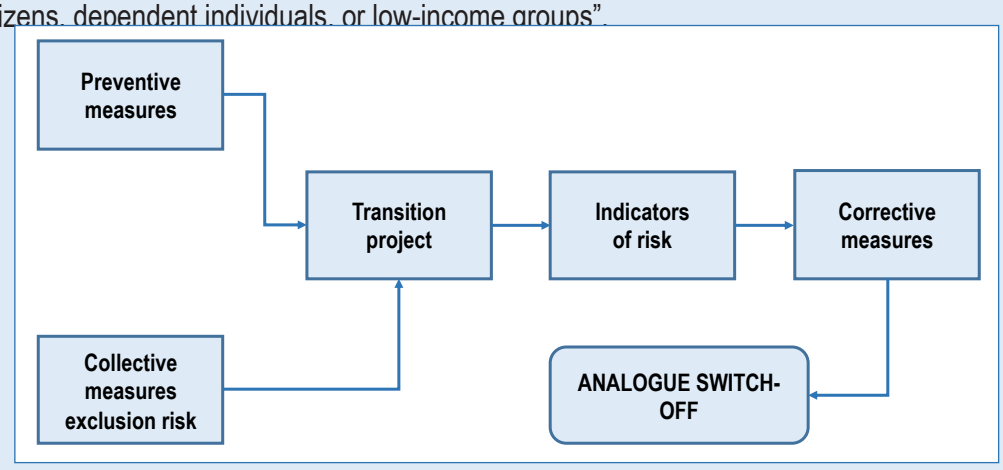
design", states the evaluation report.

An additional measure that is very useful is ***risk analysis***. It seeks to identify and assess the main problems (risks) that may arise once the Plan is launched or implemented. This assessment seeks to define, as much as possible, the preventive and corrective actions to prevent risks or to minimise their impact. The participation of the persons or bodies in charge of the intervention as well as the managers is essential when applying this methodology so the evaluator may make more grounded and feasible recommendations.

It may be a general or targeted analysis, taking into account the specific characteristics of the different areas involved in a plan when the latter has a higher degree of complexity.

Example of mechanisms for risk analysis. Evaluation of the National Plan for Transition to Digital Terrestrial Television (DTT). (AEVAL, 2009).

The report drafted by the evaluating team states that "the Plan considers the need to perform a risk analysis in order to identify the main problems that may arise once the Plan is set in motion, as well as a series of preventive and corrective actions to avoid risks or to minimise their impact. This general risk analysis is made more specific in later stages to take into account the specific characteristics of the different transition projects (TPs). Within these contingency plans, special attention is paid to certain collectives such as persons with disabilities, senior citizens, dependent individuals, or low-income groups".



3. EVALUATION CRITERIA AND QUESTIONS

Evaluation criteria in implementation evaluation

According to the Institute for the Evaluation of Public Policies and this Guide, evaluation criteria are the different points of view or approaches to the target of the evaluation, based on evidence, for its assessment. They are conditions, rules and also "principles, standards or ideas on assessment, based on which the evaluated object is assessed" (García Sánchez, 2010).

The criteria act as guidelines to formulate questions and focus the evaluation, giving it a structure that "covers the field or dimensions of a public policy or programme to be evaluated. Indeed, the questions themselves usually belong to different sets of criteria" (AEVAL, 2015).

From the point of view of implementation evaluation, reference criteria may include: suitability, coverage, coherence, complementarity, coordination, effectiveness, efficiency, implementation, participation, proportionality, sustainability, and transparency. Those that are required may be used, that is to say, only some or all, and ad hoc criteria may even be created according to the requirements of the evaluation.

Suitability

The goal of this criterion is to assess the degree of consensus between all the stakeholders, especially the managers and the target group, with regard to the development of the intervention and its instruments for coordination and participation. It is thus highly influenced by the criteria that assess these aspects and by the communication and information on the intervention which has circulated among the stakeholders.

Coverage

Coverage is defined as the degree to which a public intervention reaches a section or all of the target population. Occasionally, due to its strong connection, it has been considered a sub-criterion of implementation. The design of the intervention must be factored into this criterion, as the correct design of the target population is an essential element, which in turn is based on the correct identification and definition of the problem, which is again related to the criterion of suitability[22].

[22] Consult the Guide for Evaluating Public Policy Design (Institute for the Evaluation of Public Policies, 2020).

According to some authors, the analysis should include not only the target collective that the intervention failed to benefit but it should also study those that derived unintended benefits from it (AEVAL, 2015).

Internal coherence

This criterion assesses the relationship between the designed and implemented measures and the outputs and results. If the design evaluation seeks to make the measures and actions consistent with the proposed objectives, within a logical, formal and rational framework with a cause-effect relationship, then these characteristics must be maintained in the deployment of the intervention and in the operationalisation of the measures implemented to achieve the objectives.

Complementarity

The analysis of this criterion gauges the alignment and ad intra interactions between the measures of the intervention, localising the synergies or antagonisms, as well as the ad extra interactions between these measures and other measures or policies, thus contributing to a better solution to the problem.

Coordination

The coordination criterion seeks to assess both the structures created to optimise the necessary relationships between the managers, units, and institutions to contribute positively to the better performance of the measures and actions, and the actions performed within the framework of this coordination, collaboration or cooperation[23].

Effectiveness

Effectiveness in implementation is the degree to which the deployment actions of the intervention obtain intermediate outputs or results in line with the planned objectives, within the established deadlines.

Implementation

The criterion of implementation is the manner in which a public intervention has been implemented or launched. It focuses on analysing the structures, processes, activities, and logic developed by means of workplans and actions for implementation, in order to achieve certain results (outputs) in the target groups (AEVAL, 2015: 96).

[23] Note the different legal concept of collaboration, cooperation, and coordination for inter-administrative relationships, mentioned in Article 140 of Law 40/2015.

As mentioned throughout this Guide, there may be implementation errors that affect the fulfilment of the objectives of the evaluated intervention, or lead to unanticipated or unwanted effects. A low degree of implementation of an intervention at a certain point in time may set back the achievement of the results sought by the intervention.

Participation

This criterion assesses the role of the stakeholders during the implementation of the intervention. It includes the analysis of the mechanisms considered for this participation as well as the degree and mode of participation.

Proportionality

This criterion examines whether the implemented measures are proportionate to the effort (organisational, resource-based, etc.) made to achieve the sought results.

Sustainability

This criterion allows us to assess whether the implemented measures and their outputs can be maintained over time in order to sustain or boost the results, or if necessary, whether these measures may be redefined based on the context, resources or other factors that may change or are expected to change.

Transparency

This criterion examines stakeholder knowledge and perception regarding the adopted measures, as well as the decision-making process. In addition to dissemination, visibility and knowledge of the measures, transparency also involves the public perception of the progress made and the evolution of the problem.

Evaluation questions

The system used to create “the logical structure of the evaluation” based on a series of “queries and hypotheses that will make it possible to implement the evaluation” (AEVAL, 2010), consists of the evaluation questions, the basic unit of research in an evaluation. This logical structure includes the scope and focus of the evaluation and therefore, steers the evaluation design.

The evaluation questions constitute the operationalisation of the evaluation criteria, are framed on the basis of said criteria, and can be broken down into questions and sub-questions.

Evaluation Matrix

The ***evaluation matrix*** is the basic tool that steers the evaluation process. Normally, it consists of the evaluation questions and criteria and usually includes the indicators, sources, tools, and techniques of analysis. It seeks to make robust and credible conclusions and recommendations that are generated from the responses to said questions, according to the involved criteria.

The following section shows an example of an evaluation matrix with the common criteria for an implementation evaluation[24].

[24] The evaluation questions that are listed in the matrix are not exhaustive.

EVALUATION MATRIX

I. ¿What is the degree of implementation of the intervention as an instrument for public action, and that of each of its measures?

SUB-QUESTIONS	CRITERIA	INDICATORS	SOURCES	TOOLS
1.1. What is the level and characteristics of the deployment of the measures of the intervention?	**IMPLEMENTATION**	Number of provisions and other instruments, approved for the development of the intervention. Approved implementation structures. Degree of deployment of the intervention and its measures. Decision-making criteria in implementation structures. Degree of fulfilment of the anticipated monitoring of the intervention and its measures.	Monitoring reports. Tables and checklists for monitoring. Documents on implementation: calls, bases, etc. Guides. Statistical analysis.	Documentary analysis. Semi-structured interviews Questionnaires to beneficiaries. Interviews of managers.
1.2. Have the resources required to implement the resources been allocated?	**IMPLEMENTATION**	Budget approved. Degree of execution of spending commitment. Funds to cover the expenditure of the intervention and its measures. Personnel assigned to manage the intervention and its measures.		
1.3. Has each measure designed in the intervention been implemented, and within the scheduled deadline?	**IMPLEMENTATION**	Extent to which the measures have been set in motion Deviation with regard to the deadlines mentioned in the intervention and for each measure.		
1.4. What are the obstacles and barriers detected in the implementation of the intervention and its different measures?	**IMPLEMENTATION**	Identification of possible conflicts and detected barriers. Measures designated to solve difficulties.		
1.5. Does the intervention take into account the required tools: systems for monitoring, information collection, indicators, base line?	**IMPLEMENTATION**	Identification of the information in reliable databases. Existing databases. Number of indicators for the evaluation. Quality of the information compiled. Indicators with definition, formula for calculation and the information collection procedure.	Official information and the managers of the measures. Key stakeholders.	Documentary analysis. Semi-structured interviews of key stakeholders.
1.6. What are the measures with a level of implementation that allows us to reach the goals and objectives?	**IMPLEMENTATION**	Degree of execution awaited and confidence in the projection made by the managers. Gap between the level of implementation and the goals and measures. Analysis of coherence between inputs, activities/actions and outputs (LOGICAL FRAMEWORK)	Analysis of the results of the awaited questionnaire on execution. Statistical analysis.	Questionnaire to managers. Execution projection. Interviews of managers.
1.7. What is the assessment of the implementation of the intervention and each one of its measures by the involved stakeholders?	**IMPLEMENTATION**	Degree of agreement of the stakeholders with the implementation structures and with the results achieved.		
1.8 The units, organisations and their structures and processes are suitable for the implementation of the measures, actions and/or designed activities?	**IMPLEMENTATION**	Organisational analysis of the structures, processes, resources, etc. of the organisations or units in charge of implementing the measures.	Reports, tables and checklists for monitoring. Documents on implementation: calls, bases, etc. Guides. Statistical analysis.	Documentary analysis. Semi-structured interviews. Questionnaires to beneficiaries. Interviews of managers.

II. What is the degree of coverage and the characteristics of the population covered by the intervention?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
2.1. What is the degree of coverage in the intervention with regard to a specific target population?	**COVERAGE**	Degree of coverage of the target population.	Monitoring reports. Tables and checklists for monitoring. Documents on implementation. Microdata on the potential target population and the population that has received coverage. Surveys	Documentary analysis. Semi-structured interviews of managers.
2.2. Who have benefited from a certain intervention?	**COVERAGE**	Degree of coverage of the target population.		
2.3 What are the causes or factors that affect the coverage and participation in the programme?	**COVERAGE**	Factors or causes that affect coverage. Prevalence or probability of the different variables that affect coverage.		
2.4 Has it covered a population that should not have been covered?	**COVERAGE**	Analysis of coverage of the target population.		

III. What is the degree of coherence of the measures of the intervention with regard to the objectives sought to be achieved?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
3.1. Has the structure of the measures and outputs of the intervention been correctly defined?	**INTERNAL COHERENCE**	Degree of coherence between the measures and the requirements/problems.	Monitoring reports. Tables and checklists for monitoring. Documents on implementation. Statistical analysis.	Documentary analysis. Semi-structured interviews of managers. Questionnaire to stakeholders. Coherence analysis.
3.2. Is the intermediate result that each measure seeks to achieve, clearly identified?	**INTERNAL COHERENCE**	Indicators or definition of cross-cutting effects (synergies) of the measures. Relevance of the contributions to achieving the measures of the intervention.		
3.3. Is there any contradiction between the different measures or actions?	**INTERNAL COHERENCE**	Degree of coherence between the measures and the actions.		
3.4. Is there a deadline for achieving the measures?	**INTERNAL COHERENCE**	Evidence of the existence of measures and time-based goals in the programming documents of the intervention. Periodic monitoring is the most suitable.	Tables and checklists for monitoring. Documents on implementation.	Documentary analysis. Semi-structured interviews of managers and persons or bodies in charge.
3.5. The definition of the different measures responds to logical cause-effect relationships?	**INTERNAL COHERENCE**	Evidence of the logical definition of the causal relationships of the measures. Evidence of logical alignment in the level-based breakdown (from the strategic to the operational).	Monitoring reports. Tables and checklists for monitoring. Documents on implementation. Statistical analysis.	Documentary analysis. Semi-structured interviews of managers. Questionnaire to stakeholders. Coherence analysis.
3.6. The indicators are suitable and specific for monitoring the achievement of the proposed measures?	**INTERNAL COHERENCE**	Identification of the information in reliable databases. Quality of the information compiled.	Official information of the Administrations and the managers of the measures. Key stakeholders.	Documentary analysis. Semi-structured interviews of key stakeholders.

IV. To what extent does the intervention promote synergy between its measures?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
4.1. Are there interactions between measures that contribute to achieving the objectives?	**COMPLEMENTARITY**	% of positive interactions (synergies) between measures. % of negative interactions (antagonisms) between measures % of measures that overlap or are duplicated.	Managers, monitoring reports, tables and checklists for monitoring. Documents on implementation, guides, etc.	Documentary analysis. Semi-structured interviews of managers and involved stakeholders.
4.2. Are the objectives of the intervention related to those of other public interventions, converging towards the solution of the same problem or requirement posed?	**COMPLEMENTARITY**	No. of measures of the intervention that have features that are complementary with other public interventions. No. of studies or analyses verifying the complementary nature of the measures of the intervention with regard to other public interventions.	Managers. Guides and documents.	
4.3. Are there overlapping spheres of administrative competence between the different administrative levels?	**COMPLEMENTARITY**	% of common measures between the intervention and other public interventions in any sphere of action. % of measures that overlap or are duplicated.	Managers, reports, tables and checklists for monitoring. Documents on implementation, guides, etc.	

V. ¿What is the degree of use of the tools for coordination?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
5.1. What elements of coordination / cooperation have been included in the intervention?	**COORDINATION**	No. of tools designed and the degree of manager and stakeholder satisfaction with regard to these tools. No. of meetings / activities related to the elements of cooperation.	Managers. Monitoring reports, tables and checklists for monitoring. Documents on implementation: calls, bases, etc. Guides and documents	Documentary analysis. Semi-structured interviews of managers. Semi-structured interviews of key stakeholders.
5.2. To what extent are the measures for coordination / cooperation used in the implementation of the intervention and in each measure for every case?	**COORDINATION**	No. of meetings / activities related to the elements of cooperation and their results. Agreements reached on the basis of the mechanisms for cooperation.		
5.3. What coordination instruments have been used to reduce overlap and to maximise synergy, and what was their degree of success?	**COORDINATION**	Manage the indicators used in coordination.	Monitoring reports, tables and checklists for monitoring.	Documentary analysis. Semi-structured interviews of managers. Semi-structured interviews of key stakeholders.
5.4. Have mechanisms for communication, information exchange, etc. with these bodies been established? Of what type? What is their scope?	**COORDINATION**	Agreements reached on the basis of the mechanisms for coordination.	Documents on implementation, calls, bases,... Reports, tables and checklists for monitoring.	Semi-structured interviews.

VI. What is the degree of knowledge regarding the intervention and the involvement of interested stakeholders?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
6.1. What is the degree of stakeholder participation (institutions, society and experts) in identifying and solving problems?	**PARTICIPATION**	Information and training to promote greater public participation and mechanisms for participation among all the stakeholders.	Systems for public participation in the drafting and implementation of the intervention.	Documentary analysis. Survey or qualitative interview or group discussion with stakeholders and interview of the persons or bodies in charge of implementing the policy.
6.2. Has the intervention been implemented in participation with the social agents and relevant institutions, taking into account their suggestions?	**PARTICIPATION**	Types of mechanisms for participation.	Degree of pluralism of the participating group of stakeholders.	Map of stakeholders.

VII. To what extent have the actions of the intervention obtained results aimed at achieving the planned objectives?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
7.1. Have the instruments used and their outputs led to the awaited results or objectives?	**EFFECTIVENESS**	Analysis of the instruments for the information system and the statistical analysis of the indicators and their evolution. Analysis of the implementation and results of the process with suitable indicators.	Monitoring reports, tables and checklists for monitoring. Documents on implementation: calls, bases, etc. Guides and documents	Documentary analysis. Interviews. Microdata analysis. Internal management documents.
7.2 Does the level of implementation of the intervention let us obtain the intermediate results defined, and within the planned deadline?	**EFFECTIVENESS**	Degree of fulfilment of the plan for executing the measures. Degree of fulfilment of the intermediate goals Degree of confidence in the projected compliance with the execution goals and deadlines		
7.2. To what extent have the measures adopted in the intervention contributed or are contributing to solving the problem?	**EFFECTIVENESS**	Analysis of the information system documents.		

VIII. To what extent have the resources used been suitable for the implemented measures?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
8.1 To what extent has the use of resources to implement the measures been sufficient?	**PROPORTIONALITY**	Analysis of the implementation and results of the process with suitable indicators.	Monitoring reports, tables and checklists for monitoring.	Documentary analysis. Interviews. Microdata analysis. Internal management documents.
8.2 What is the budgetary outlay for human resources in the implementation?	**PROPORTIONALITY**	Analysis of the planning documents and the information and monitoring system.		

IX. To what extent can the measures and their outputs be maintained over time?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
9.1 Have the resources allocated to the intervention undergone any modification over the course of the implementation? How do they influence the deployment instruments?	**SUSTAINABILITY**	Analysis of the implementation and prediction models.	Documents and monitoring reports	Documentary analysis.
9.2 Is there any legislative change that may interfere with the implementation?	**SUSTAINABILITY**	Analysis of the implementation and prediction models.	Legislation documents	Documentary analysis.
9.3 To what extent do legislative, environmental, and other changes, affect the implementation in the medium and long term?	**SUSTAINABILITY**	Analysis of the implementation and prediction models.	Monitoring reports, tables and checklists for monitoring.	Monitoring reports, tables and checklists for monitoring.

X. ¿What is the degree of awareness of the intervention among the stakeholders and the general population?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
10.1 What is the degree of perception of the stakeholders regarding the adopted measures?	**TRANSPARENCY**	Information and training to promote greater public participation and mechanisms for participation among the stakeholders.	Monitoring reports, tables and checklists for monitoring.	Survey or qualitative interview or group discussion to/with stakeholders and interview of the persons or bodies in charge of implementing the policy.
10.2 What is the degree of the citizens' knowledge regarding the implemented measures?	**TRANSPARENCY**	Information and training to promote greater public participation and transparency regarding the Administration's actions.		Survey or qualitative interview or group discussion to/with citizens, regardless of whether they belong to the target group or not.
10.3 Have all the proposed means of dissemination been used? To what degree?	**TRANSPARENCY**	Analysis of the implementation and results of the process with suitable indicators.		Monitoring reports, tables and checklists for monitoring.

XI. To what extent is there consensus on the development of the intervention, with regard to the instruments for coordination and participation?				
SUB-QUESTIONS	**CRITERIA**	**INDICATORS**	**SOURCES**	**TOOLS**
11.1 To what extent is there consensus on the development of the intervention, with regard to the instruments for coordination and participation between managers and the target population?	**SUITABILITY**	% of satisfaction with the instruments for coordination and participation used among different stakeholders.	Monitoring reports, tables and checklists for monitoring.	Survey or qualitative interview or group discussion to/with the target population and interview of the persons or bodies in charge of implementing the policy.
11.2 To what extent is there a greater deviation in perception on the development of the intervention, with regard to the instruments for coordination and participation between managers and the target population?	**SUITABILITY**	% of satisfaction with the instruments for coordination and participation used among different stakeholders.	Monitoring reports, tables and checklists for monitoring.	

4. ANALYSIS TECHNIQUES IN IMPLEMENTATION EVALUATION

For the implementation evaluation, there are different tools and techniques that allow the evaluator to obtain rigorous proof that responds to the evaluation questions or to analyse the different questions mentioned in the evaluation.

This Guide provides a brief description of the most relevant social research techniques that are of the greatest use and validity for evaluation in general. The most traditional classification of available techniques is that which distinguishes between qualitative and quantitative techniques.

Thus, among ***qualitative*** techniques, we have documentary analysis, interviews, discussion groups, nominal group techniques, discourse analysis, SWOT analysis and case studies. And among ***quantitative*** techniques we have purely descriptive statistical analyses, statistical inference or relations between the variables or phenomenon under study, either by means of statistical association or more complex analyses, such as simple linear regression models, multiple linear regression models, logistic regressions, etc.

When analysing the implementation, qualitative methods allow us to obtain in-depth information on the perceptions and opinions of a group of persons on a certain question.

These methods are normally supplemented with quantitative methods as they arise from the questions that the implementation evaluation seeks to answer, and which are quantitative in nature, such as the execution of the measures or the resources used. The latter, nevertheless, are mainly used in the evaluation of the results.

Figure 19. Analysis techniques in an evaluation. Source: Author's own.

Type of technique		Purpose/nature
Qualitative techniques	Documentary analysis	Exploratory. Applicable at any stage.
	Interviews	Exploratory. Applicable at any stage.
	Group discussions	Collecting qualitative information. Facilitating comprehension, credibility and acceptance.
	Nominal Group Techniques (NGT)	Structured analysis of ideas and problems.
	Discourse analysis	Analysing all discourses and the contexts in which they are produced.
	SWOT	Reducing uncertainty and define strategies.
	Case studies	Analysis of results and impacts.
Quantitative techniques	Survey	Obtaining descriptive information or other type of information in order to apply other techniques.
	Linear regressions	Analysis of explanatory causes and estimating effects.
	Logistic or probabilistic regressions	Analysis of explanatory causes and estimating effects.
	Cost-benefit analysis	Knowledge of differentiated impacts. Efficiency analysis.
	Cost-effectiveness analysis	Effectiveness analysis based on a relevant criterion.
	ARIMA Models	Time-series analysis.
	Multi-level analysis	Studying contextual factors, either by hierarchy or by levels.
	Stochastic frontier models	Measuring efficiency in terms of input maximisation.
	Factorial analysis	Reducing underlying dimensions.
	Impact evaluation methods	Measuring net effects attributable to a public intervention.
Mixed	Multiple criteria analysis	Structuring and combining assessments taken into account in a decision.

Qualitative techniques

Documentary analysis

The documentation associated with the intervention is a major source of information. It refers to the documents of the intervention, programming, applicable legislation, internal orders, guidelines, budget justifications, monitoring and results reports. Basically, everything that includes the institutional point of view. (Chen 1990: 66). It is also interesting to perform a comparative analysis of the current literature and evaluations on the topic.

Interview

According to Dezin and Lincoln (2005), the interview is "a conversation, it is the art of asking questions and listening to the answers". This definition is based on a simple relationship between the researcher and the interviewee where the researcher asks questions that may

range from opinion surveys or questionnaires, that is to say, highly structured instruments, to open interviews where the researcher may even be questioned or queried by the interviewee. In qualitative research, the interview is not necessarily based on closed and structured questionnaires but on the contrary the researcher may repeat these meetings until all emerging or relevant topics have been clarified.

There are different types of interviews:

Structured interviews

In this type of interview, the questions to be asked are previously planned. A targeted and sequential list of questions is prepared. The interviewee cannot make comments or appraisals. These are closed questions; therefore the answers must be specific and exact.

Semi-structured interviews

The researcher prepares the questions beforehand on the basis of a thematic script. The questions shall be open and in contrast to structured interviews, the interviewee may express their opinions, qualify their responses, and even deviate from the initial script.
These are the most commonly used interviews in all types of evaluation, including implementation evaluations.

Figure 20. Example of semi-structured questions in the Evaluation of the measures for streamlining and improving the management of Temporary Disability. Source: (AEVAL, 2009).

Evaluation questions	Evaluation criteria
What part of the evolution of the expenditure cannot be explained by the working population, the regulatory base, or by ageing?	Suitability
Have the General State Administration agencies been equipped with the organisational instruments?	Coherence
Is there complementarity and coordination of the intervention between the different entities responsible for managing temporary disability due to common contingencies and has the coordination been effective?	Complementarity and effectiveness

Unstructured or open interviews

These are generally known as in-depth interviews. In this case, the objective is “to understand the interviewees’ perspectives with regard to their lives, experiences or situations, expressed in their own words” (Taylor and Bogdan, 2008). These interviews are modelled after a conversation between peers and not a formal exchange of questions and answers. They require multiple meetings with the interviewees. There are three types of in-depth interviews: life histories, learning about events and activities that cannot be observed directly, and interviewing an extensive group. These three are of great use in applied social research but not directly in evaluation, as their goals are different.

Group discussion

Group discussion is a qualitative technique which brings together a group of people to obtain information on a specific topic, conducted by an interviewer.

Group discussion is a highly valuable technique to obtain information or qualitative evidence, as it generates a series of interactions among the people who are part of the group and it aids in obtaining information that is different from what is obtained in individual interviews.

When organising a group discussion, it is very important to be clear about the objective that is sought.

Based on each case, a group discussion may have different objectives:

- To share information and knowledge.
- To provide different perspectives.
- To find a common denominator.
- To come to an agreement.
- To compile qualitative information on perceptions, motivations, opinions, attitudes, etc.

There are different stages of development of a group discussion:

- Establishing objectives: The first step is to set the group objective and based on this decision, define the type of group (more open or more closed) to be formed, the participants to be invited (the sample), and develop tools for the group's functioning (script, schedule, activities, etc.).

- Selecting participants: In this stage we shall define the characteristics of all the participants and select the persons invited to form part of the group discussion. It is termed an "international" sample as it is not extracted on the basis of statistical criteria, nor is it a random selection, rather people are selected on the basis of their relationship with the topic under study.

- Preparing the group discussion: In this third stage, the group is planned, both with regard to the questions to be asked or the activities to be performed, and the logistic aspects.

- Group organisation: The group discussion is constituted.

- Analysing information and drawing conclusions: In this last stage, conclusions are obtained from the observations and results of the group work.

There are different types of group discussions, depending essentially on the role adopted by the group moderator and the level of conducting; from very open groups where different members of the group participate in a debate on the basis of pre-set questions, to other more focused ones that apply specific group dynamics techniques and lead the group participants towards a concrete point.

Generally, group discussions fall into two large classes:

Focus group

It is a group session, conducted by a moderator. It consists of a debate between different persons based on a list of questions that have been defined in advance and where the moderator suggests issues or asks questions and the group participants respond to them.
The goal of this technique is to obtain in-depth information on a specific topic by listening to a group of persons related to the topic under analysis.

Group dynamics

In this case, we are dealing with a programmed session with a series of activities and specific group dynamics that seek concrete objectives.

The objective of this technique is highly varied, although it focuses on analysing and diagnosing, or seeking symptoms and requirements of the analysed situation. Its goal is to propose alternatives and analyse the current situation with regard to certain envisaged objectives.

This technique has its advantages and disadvantages, as it helps to pool ideas, share experiences, and build consensus. It also helps to find the common denominator between the participants. Conversely, it may lead to organisation and logistics problems and it requires prior experience. Other disadvantages are that there may arise problems, arguments, and complaints that the moderator may not be able to control.

Nominal Group Technique

Nominal Group Technique (NGT) is a creative technique for analysing problems that combines individual opinions and facilitates the decision-making process. It helps to identify the elements of a situation or problem, gives partial or total solutions to them, and establishes priorities by consulting a group of persons while respecting their anonymity.
Its development consists of five stages:

- Formulation stage. In the first stage, questions linked to the problems, obstacles, or difficulties are posed.

- Reflection stage. In the second stage, all participants are asked to reflect on these questions silently and individually.

- Grouping alternatives stage. In the third stage, aided by the group participants, the researcher groups all the reflections made in the first and second stages, according to the degree of similarity of each, as judged by the group.

- Debate stage. In the fourth stage, a debate is initiated on the importance of each question that has been posed. The group votes on the groups of ideas.

- Voting stage. The fifth stage corresponds to the hierarchical arrangement of the alternatives.
  The process concludes with the final report drafted by the expert, who passes it on to the relevant individual or body so that they adopt the required measures and attempt to solve the problems or questions posed in the NGT, or take into account the suggestions made by the participants.

This technique has a series of advantages, among them the systematic and orderly analysis of problems, as well as highlighting proposals for decision-making by combining individual creative responses that become qualified group opinions.

When applied to public policies, this technique allows us to identify problems and their areas of improvement. It also lets us analyse their causes and solutions. In the Evaluation of the Human Resources Quality Plan of the General Services of the Administration of the Autonomous Regions of the Balearic Islands (ACAIB) (AEVAL, 2015), this technique was applied to three groups:

- The first group consisted of nine HR managers of the General Secretariats of the council offices of the Autonomous Region of the Balearic Islands (or CAIB in Spanish).

- The second group consisted of eleven heads of all the CAIB councils with a common denominator, they had staff members and at least four years of experience in public administration.

- The last group consisted of ten ACAIB civil servants. This group was characterised by its heterogeneity.

The methodology used was common to all three groups and it unfolded in the following manner:

1. Presenting the participants.
2. Formulating the first question. In your opinion, what are the main problems that affect the management of ACAIB personnel?
3. Silent generation of ideas.
4. Collecting the ideas-responses.
5. Group discussion of the ideas-responses, interpretation, and clarification.
6. Voting.
7. Break.
8. Formulating the second question: In your opinion, how can the management of ACAIB personnel be improved?
9. Session end.

Delphi Method

It is a group technique that allows us to classify expert opinions by means of an interactive process of individual questions.

It consists of four successive rounds of questionnaires. The responses are summarised in order to draft the next consultation and an agreement is reached.

After the first questionnaire round, we come to the next stage where the experts must again respond in view of the results of the first questionnaire and justify their differences with regard to the group. In the third stage, the expert is asked to comment on the arguments that deviate from the majority, and in the last stage, a final consensus is reached. The following section displays a brief overview of the process:

- Early stage: Defining objectives, identifying interviewees, and selecting the areas of study.
- Development stage: Designing and drafting the first questionnaire. Process and obtain the average of all the results. Identifying points of divergence and homogeneity. The results of the first questionnaire are used to draft the second questionnaire and so on and so forth.

Figure 21. Delphi Questionnaires. Source: Methodological Guide on Auditing for Inspectors of the General State Administration Services (December 2009).

	QUESTIONNAIRES		
	1	2	3
CONTENT	1 or 2 open questions	Transmit, prioritise and comment in favour of / against	Transmit and revise priorities
ANALYSIS	Classify, summarise and quantify	Identify areas of agreement / disagreement and establish priorities	Establish final results

- Final stage: The results are analysed and the conclusions drafted.

The benefit of this technique lies mainly in the insistence generated by presenting the same questionnaire several times. That is to say, the results of the previous questionnaires help experts to progressively learn about the different points of view so they may continue to modify their opinion if the arguments presented appear to be more suitable than their own.

SWOT Analysis

A SWOT analysis is a simple and general tool for taking strategic decisions. The main goal is to help find strategic elements and use them to make changes in the organisation by consolidating strengths, minimising weaknesses, taking advantage of opportunities, and eliminating or reducing threats.

It derives its name from the initials S (for strengths), W (for weaknesses), O (for opportunities) and T (for threats).

This technique is based mainly on two types of analysis, internal and external.
In an internal analysis, the objective is to detect the weaknesses and strengths of the organisation: to remedy the first and to promote the second. Different aspects are studied for this purpose: production, organisation, human or personnel resources and finances.

External analysis focuses mainly on detecting threats and opportunities. For this we shall consider the environment of the organisation, interest groups, legislative, demographic, and political issues. These points are very revealing when it comes to defining strategies that seek to combat threats and take advantage of opportunities.

Once the strengths, weaknesses, opportunities, and threats have been identified, the SWOT Matrix may be created, which allows us to visualise and summarise the current situation of the organisation. With the results of the SWOT analysis, a strategy must be defined.
Below is an example of a SWOT matrix:

Figure 22. SWOT Matrix. Source: Author's own.

		STRENGTHS (S)	WEAKNESSES (W)
		Of the organisation	
OPPORTUNITIES (O)	Of the environment	Take advantage of the opportunities offered by the environment, using the organisation's strengths.	Take advantage of the opportunities offered by the environment, overcoming the organisation's weaknesses.
THREATS (T)		Use the organisation's strengths to avoid the threats posed by the environment.	By reducing the weaknesses of the organisation, we avoid threats.

The advantages of this technique are mainly that it leads to an awareness of existing problems, their characteristics and how they interact with the context, the organisation, or the institutional framework, as well as the risks and opportunities generated by the environment that surrounds said organisation.

Quantitative techniques

Survey

A survey is one of the most frequently-used techniques in any type of evaluation, including design evaluation as it allows us to clearly identify design problems as well as the perception of the stakeholders and the existing difficulties from the point of view of the managers, stakeholders or the targets of the intervention. It also allows us to obtain results from a specific territory that may be generalised to the entire population. As a source of primary data, it allows the evaluator to arrange them in the most convenient way possible to obtain the necessary information for the research.

It is a reliable but expensive technique and requires an exhaustive knowledge of the intervention and a thorough preparation of the framework of analysis by the evaluator.

When performing a survey, the first step is the ***sample selection***, which must be as representative as possible of the reference population, in order to make generalisations with regard to the population. Random sampling methods ensure the best sample representation. This means that any individual in the selected sample has the same probability of being selected.

Another aspect to be taken into account to optimise the results of the survey is the selection of the sample size. This requires a considerable knowledge of sampling techniques, a topic which is beyond the scope of this Guide. Nevertheless, it must be remembered that the greater the sample size, the lower the estimation error and thus, the more significant the results, although not indefinitely.

On other occasions, when the total population is not excessively high, all the members may be surveyed. Let us take, for example, a survey of organisations or units numbering between 100 and 200.

Once the sample size is selected, we come to the ***survey design***, which is the instrument for compiling and measuring data, and is characterised by a series of questions arranged according to a specific logic. Its design must be adjusted to the established objective and for this, we must be clear about what we wish to ask and above all, how to ask: It is important for the questions to be clear and concise, and flexible and "comfortable" answers must be provided to the interviewee. Finally, the questionnaire must not be very long.

There are different types of questions: ***open, closed, semi-open (or semi-closed)***.

With regard to the ***mode of administration*** of the questionnaires, they may be self-administered, in-person, telephone, postal or online surveys. The decision to opt for one or another depends on the advantages and disadvantages of each of them according to the topic under study, the available time and financial resources and the target population of the survey.

In-person surveys are most frequently used in social research. They have the advantage of a more complete form of obtaining information and allow researchers to capture the environment surrounding the survey. But it has the disadvantage of being expensive, slow, and difficult to access by certain populations.

- The main requirement for telephone surveys is that the surveyor must have a comfortable format. When drafting the questionnaire it is important to assess whether the design, duration, order, and interpretation are the most suitable. Currently they are mostly performed as Computer Assisted Telephone Interviewing (CATI), which lowers costs and the time required to perform them. However, it is not appropriate for delicate topics or complex questions. This survey mode may suffer from technical errors.

- In a postal or online survey, the interviewee reads the questionnaire and notes down their responses. There is no interviewer and therefore, a letter of presentation is required. It is a cost-effective technique and requires few personnel to perform the survey. It gives anonymity and flexibility of time to the interviewee. Its disadvantages include low levels of response and errors in filling out the questionnaire.

With regard to ***specific types of surveys***, we may mention:

- Omnibus surveys that allow us to include various topics, research or evaluation goals in a single survey. It is cost-effective, as instead of multiple surveys, only one is performed, thus sharing the research costs, and formulating a reduced number of questions in the same questionnaire and targeting the same sample. This type of survey is generally meant for large populations to achieve a financially feasible study. The questionnaire follows the same criteria as the interview but distinguishes itself by being arranged into different sub-questionnaires or modules with regard to different topics or outputs.

- Panel survey is a quantitative marketing research technique that is performed periodically on the same representative sample of a specific population.

Descriptive statistics

Once the survey is performed and the data has been filtered, they are analysed by means of ***descriptive statistical techniques***.

- Absolute and relative frequencies (the number of times an event is repeated and what it represents at the level of the population, respectively).

- Measures of centralisation are used (mean, median and mode) to obtain an overview of the data.

- Measures of dispersion, that provide an idea of variation in the sample data. They are useful when assessing the reliability of measures of centralisation such as the mean. They have an inverse relationship, the higher the measure of dispersion, the lower the representativeness of the measure of centralisation. The most well-known are variance and range; the range measures the difference between the maximum and minimum value that the observations can reach; variance measures the distance between the data and the mean.

These descriptive statistical techniques are characterised by their study of random phenomena; therefore their results are not precise and are accompanied by a certain degree of uncertainty. To measure this degree of uncertainty, we use statistical inference techniques.

Statistical inference

Statistical inference techniques give us the answers to questions such as: What variables influence the incident? How do the variables influence the incident? Is it possible to obtain a model that explains the incident and allows us to predict its behaviour? Some of these techniques are described below.

Lineal regression

In didactic terms, regressions seek to explain a variable or phenomenon that is deemed independent or endogenous by means of a series of facts, phenomena or variables that are called regressors, covariates or explanatory factors. It is the latter that may explain to a certain degree a phenomenon, behaviour, or reality.

Regression allows us to adjust a point cloud to a function where the endogenous or independent variable is explained partially through regressors or dependent variables, at the same time that the contribution of each dependent variable to the aforementioned explanation is determined. The difference between the real values and the explanation of the endogenous variable by the regressors is what constitutes the error term or random term.

When the independent variable is continuous and the function that links the endogenous variable with the regressors is linear, it is called linear regression. Apart from this configuration element, the assumptions on which the adjustment is made are: non-correlated regressors, their variance is constant (homoscedasticity), the errors in the measurement of each are inter-related and add to the total error, and the expected value is equal to zero, that is to say, the errors of a similar magnitude and opposite signs are equiprobable.

Provided the target of the analysis permits it and there is sufficient high-quality data, this technique can provide useful evidence for an evaluation.

Example: Evaluation of the Plan for Measures to Improve Cross-Border Healthcare Services (AEVAL, 2013).

"The third step is to apply the personnel estimation model. For this, a linear regression model has been developed that estimates the staffing of each service that would correspond to its calculated complexity, and identifies the services that exceed or fall short of said estimate. The dependent variable considered when building the model is the total occupied personnel on 31 December 2012, and as sub-group, the inspectors (both A1 and A2). The independent

variables or predictors are the total complexity of the services and the total number of entries (records) in groups of a thousand. Additionally, dummy variables are created for the qualitative variables of time and service so that they are considered when calculating the estimate. Of the models built, the one with the best statistical adjustment has been selected".

Logistic or ordinal regression

Linear regression is a regression module where the variable or fact to be explained takes either two values (the phenomenon takes place or it doesn't, i.e., yes or no) or very few values (for example a scale of 5 values that measures intensity as a lot, enough, little, or nothing). Or to put it in another way, the variable to be explained is not continuous or the function is logistic. Similar to linear regression, logistic regression allows us to adjust a cloud of points to a function where an endogenous variable is partially explained through regressors.

Cost-benefit and cost-effectiveness analyses

Before assigning monetary resources to a public or private intervention, the quotient of discounted cash flows between the allocation of resources (cost) and their returns (profits) allows us to assess in absolute terms the convenience of allocating said resources or eventually of allocating them to alternative options. Occasionally, when the costs of the evaluated event are not explicit owing to the fact there is no market that reveals them, the so-called shadow prices are adopted as prices that they would have under perfectly competitive conditions.

The cost-effectiveness analysis is a variant of cost-benefit analysis that is applied when there is a lack of prices to assess the objective or set of objectives that the intervention seeks to achieve. To this end, cost would be that which allows the maximisation of the objective. When alternative interventions to achieve the same objective are compared, the selection criteria shall be to consider the intervention that helps to reach the objective at a lower cost, and at equal costs helps to maximise the objective.

Whenever faced with a problem that is resolved by cost-benefit or cost-effectiveness analysis, these techniques may constitute evaluation criteria.

ARIMA Model

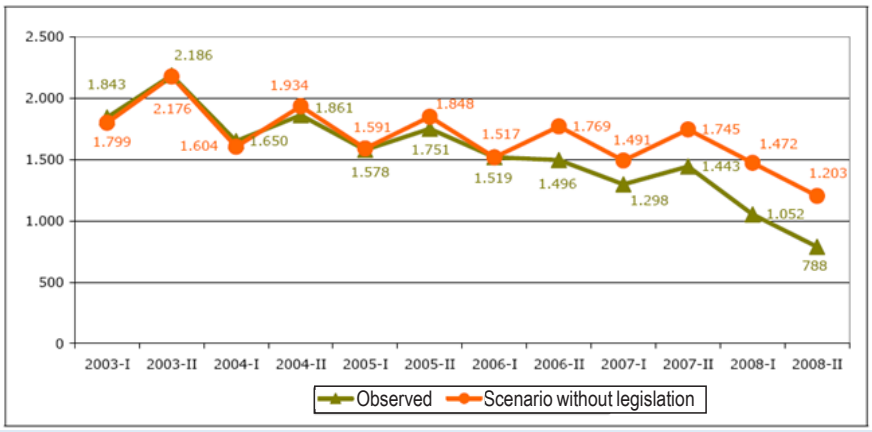
The ARIMA model (Autoregressive Integrated Moving Average) is a technique used to establish patterns of behaviour or of facts with the goal of forecasting. It does not use other variables as in regression techniques, but past data or values. Each observation is determined by earlier values in time. The ARIMA model (p, d, q) is denoted by means of three parameters - p, d, q, non-negative integers - that highlight the order of the three parts of the model: autoregression, integration and moving-average.

ARIMA models are used in evaluation to define patterns and make predictions. It is a dynamic time-series model, that is to say, future estimates are explained by the data of the past and not by independent variables.

Example: Evaluation of the Strategic Plan for Road Safety 2005-2008 (AEVAL, 2009).

This evaluation uses an ARIMA model to study the impact of certain key variables on the victims' time series, primarily the legislative changes generated by the Plan (points-based driver's license and reforming the Criminal Code, above all). The study highlights that although there was already an underlying cause that implied a descent in the number of fatalities, what is certain is that "the impact of the plan and especially, of the plan put into motion from 2006 (especially the points-based license and the Criminal Code reform) has been responsible for reducing almost all fatalities over 24 hours". The following figure displays the differences between the observed situation (green line) and that which would have occurred without the implementation of the measures according to the ARIMA model (orange line).

Total number of highway fatalities over 24 hours



Multi-level analysis

Multi-level models of analysis (hierarchical linear models, linear mixed-effect and nested models, among others) are models with parameters that vary in more than one dimension. They are of use when discerning what part of an effect may be attributed to one cause and what part to another, when both are present at the same time.

For example, in research on education, they would be useful to measure what part of the students' performance is due to the teaching method or to the school or institution where they study, and what part to other variables such as the social background of the students.

Frontier or efficiency models

Another tool that helps us to analyse certain phenomena in terms of efficiency or inefficiency of the resources used with regard to the maximum potential results that may be obtained with them. These are frontier analyses of the production or cost function. Based on the definition of a Production–Possibility Frontier (PPF), these models display, firstly, the parameters that define the frontier by their functional characterisation and subsequently, the efficient options (the ones that are situated on the production frontier) and the inefficient ones. An allocation of economic resources is efficient when it is situated on the PPF.

There are three types of frontier models, of which essentially two are important:

- Nonparametric or mathematical models. This is data envelopment analysis (DEA). It uses mathematical programming to establish the set of observations that estimates the frontier and which do not require a previous functional form.

- Parametric or stochastic frontier models. It allows the estimation of the frontier functional form, costs, or benefits, the parameters, and its advantage is that it incorporates the specification error and allows us to distinguish the effects of noise or inefficiency error.

Stochastic frontier models are included in evaluation as an analytical option for applying the efficiency criteria.

Factorial analysis and principal component analysis (PCA)

When faced with a high number of variables with different degrees of correlation or linear dependency between them, both techniques may be used to reduce them to a set of factors or components that provide a synthesis of the phenomenon under study. Principal component analysis and factorial analysis both reduce the number of explicative variables, but differ in how they do it.

In the case of factorial analysis, the original variables are grouped by factors, so that they may be defined as linear combinations of the factors and explain the covariance or correlations between them.

Conversely, principal component analysis (PCA) defines new variables or linear independent components from the original variables. By means of a linear transformation, it defines a new system of coordinates for the original dataset where the highest variance is assigned to the first principal component, the second highest variance to the second component and so on, until the total variance contained in the original variables is saturated. In PCA, components are calculated as linear combinations of the original variables, normally after centring the data in the average of each.

Both techniques may be used in evaluation for exploratory, analytical, or confirmatory purposes.

Impact evaluation methods

Impact analysis or evaluation methods allow us to determine what part of the observed effects or results of a phenomenon are solely and exclusively attributable to a fact, in this case, a programme or an intervention. They are also called counterfactual methods.

The advantage of this type of methods is that they statistically isolate multicausality and isolate the effects, so that it may be stated with statistical rigour that the observed results are the result of a factor, fact, programme, or intervention.

Impact evaluation methods compare the results observed in the population, drawing a distinction between the target group or persons who receive an intervention and those who do not receive it, called the control group. If both groups are statistically similar or identical, the observed result can only be dependent on the treatment.

Impact evaluation tools or techniques may be divided into those based on experimental models, when it is possible to define in advance the phenomenon that receives or does not receive the intervention, through random processes; and quasi-experimental models where it is not possible to randomise in advance.

Multiple criteria analysis

Occasionally, the target of the evaluation may be assessed according to various criteria. On the basis of the weight of each criterion and according to a ratings scale, it is possible to quantitatively measure the joint application of different criteria and to sum them up in a number (the sum of the outputs: the weighting applied to the criterion by points attributed to the criterion), and thus compare alternatives.

LIST OF FIGURES

WORKS CITED

AEVAL. (2009). Evaluación de las medidas de racionalización y mejora de la gestión de la Incapacidad Temporal. Madrid.

AEVAL. (2009). Evaluación del Plan Estratégico de Seguridad Vial 2005-2008. Madrid.

AEVAL. (2009). Evaluación del Plan Nacional de Transición a la Televisión Digital Terrestre (TDT). Madrid.

AEVAL. (2010). Fundamentos de evaluación de políticas públicas. Madrid. Obtenido de http://www.aeval.es/export/sites/aeval/comun/pdf/evaluaciones/Fundamentos_de_evaluacion.pdf
AEVAL. (2011). Evaluación del Plan Español de Energías Renovables. Madrid.

AEVAL. (2011). Evaluación del Programa de Ayudas a las Actuaciones de Reindustrialización (REINDUS). Madrid.

AEVAL. (2014). Evaluación del Plan de Ahorro y eficiencia energética 2008-2012. Madrid.

AEVAL. (2014). Evaluación del Plan de Medidas para la Mejora de los Servicios de Sanidad Exterior. Madrid.

AEVAL. (2015). Evaluación del Plan de Calidad de los Recursos Humanos de los servicios generales de la Administración de la Comunidad Autónoma de las Illes Balears. Madrid.

AEVAL. (2015). Guía práctica para el diseño y la realización de evaluaciones de políticas públicas - Enfoque AEVAL. Madrid.

AEVAL. (2016). Seguimiento de la formación dirigida a los empleados públicos: Administración General del Estado, organizaciones sindicales, Entidades Locales y Comunidades Autónomas. Madrid.

AEVAL. (2017). Evaluación intermedia del Plan Nacional de Desarrollo Rural (PNDR). Madrid.

Aguilar, L. (1993). Problemas públicos y agenda de Gobierno. En L. Aguilar, El estudio de las políticas públicas. Antologías. México: Miguel Ángel Porrúa Editores.

Alvira Martín, F. (1991). Metodología de evaluación de programas. CIS.

Alvira Martín, F. (2002). Manual para la elaboración y evaluación de programas de prevención de abuso de drogas. Madrid: Comunidad de Madrid, Agencia Antidroga.

Alvira Martín, F. (febrero de 2018). Formas y niveles de participación. Material didáctico, INAP. Madrid.

Alvira Martín, F. (marzo de 2018). La evaluación de la participación política directa. Material didáctico, INAP. Madrid.

Bardach, E. (1993). Problemas de la definición de problemas en el análisis de políticas públicas. En L. E. Aguilar, Problemas publicos y agenda de gobierno. México: M.A. Porrúa Editores.

Birkland, T. (1997). AfterDisaster: Agenda-Setting, Public Policy and Focusing Events. Georgetown University Press.

Bueno Suárez, C., & Osuna Bañeza, J. L. (2013). Reflexiones epistemológicas y metodológicas para la evaluación de políticas públicas. Andamios. Revista de investigación social.

Cardozo, M. (2008). Gestión y evaluación participativas en políticas sociales. Política y Cultura(30), 137-163.

CEPAL. (2005). Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas.

Chen, H. (1990). Theory-driven evaluations. New Park, CA: Sage Publications Inc.

Chen, H. (1994). Theory-driven evaluations: Need, difficulties and options. Evaluation Practice, 15(1), 79-82.

Chen, H.-T., & Rossi, P. H. (2016). Evaluación con sentido: el enfoque basado en la teoría. En C. Maldonado, & G. Pérez, Antología sobre evaluación. La construcción de una disciplina (págs. 51-67). México D.F.: Centro de Investigación y Docencia Economías: Centro CLEAR para América Latina.

Comisión Europea. (s.f.). El nuevo periodo de programación 2000-2006: Documentos de trabajo metodológicos. Documento de trabajo 3. Indicadores de seguimiento y evaluación: Orientaciones metodológicas. Obtenido de https://ec.europa.eu/regional_policy/sources/docoffic/working/doc/indic_es.pdf

Comisión Europea. (s.f.). El nuevo periodo de programación 200-2006: Documentos de trabajo metodológicos. Documento de trabajo 3. Indicadores de seguimiento y evaluación:

Orientaciones metodológicas. Obtenido de https://ec.europa.eu/regional_policy/sources/docoffic/working/doc/indic_es.pdf

Comisión Europea. EuropeAid - Oficina de Cooperación. (Marzo de 2001). Manual Gestión del ciclo del proyecto.

Denzin y Lincoln. (2005). The SAGE Handbook of Qualitative Research.

Departamento de Administración Pública y Justicia. Gobierno Vasco. (2014). Modelo para la evaluación de procesos participativos en las administraciones públicas. Un navegador para acompañar el viaje. Un dispositivo para aprender del viaje.

Díez López, M. (mayo-diciembre de 2004). La evaluación basada en la teoría y su aplicación a la política regional: algunas reflexiones teóricas. gapp(30-31), 57-70.

Dunn, W. (1981). Structuring policy problems. Public Policy Analysis Fih Edition. Pearson.
Feinstein, O., Ligero Lasa, J. A., Rein, M., Schon, D., Majone, G., Weiss, C. H., . . . Picciotto,

R. (2016). La evaluación de políticas. Fundamentos conceptuales y analíticos. Argentina: Corporación Andina de Fomento. Serie Estado, Gestión Pública y Desarrollo en América Latina.

García Sánchez, I. (2010). Sistema de evaluación. Obtenido de http://www.eumed.net/libros/2010b(687/

Grau-Solés, M., Íñiguez-Rueda, L., & Subirats, J. (2011). ¿Cómo gobernar la complejidad? Invitación a una gobernanza urbana híbrida y relacional. Athenea Digital(11), 63-84.

Greene, J. (2016). La evaluación como defensa. En C. Maldonado, & G. Pérez, Antología sobre evaluación. La construcción de una disciplina (págs. 106-119). CIDE.

Hoppe, R. (2018). Rules of thumb for problem-structuring policy design. Policy Design and Practice.

Instituto para la Evaluación de Políticas Públicas. (2019). El diseño de las intervenciones públicas. Guía de evaluación. Madrid.

Instituto para la Evaluación de Políticas Públicas. (2019). La evaluabilidad de las intervenciones públicas. Guía de evaluación. Madrid.

Ivalùa. (2009). Guía práctica 3. Evaluación del diseño.

Ligero Lasa, J. A. (2016). Dos métodos de evaluación: criterios y teoría del programa. En La evaluación de políticas. Fundamentos conceptuales y analíticos.

Lima Facio, & Aguilar Astorga. (Abril de 2011). La importancia del diagnóstico en las políticas públicas. Contribuciones a las Ciencias Sociales. Obtenido de ww.eumed.net/rev/cccss/12

Linder, S., & Peters, B. (1984). From social theory to policy design. Policy formulationa and the challenge of conscious design.

Macía, M. (2017). Metaevaluación de la dimensión política en las evaluaciones de AEVAL. Cinco experiencias de evaluación. Trabajo de Fin de Máster, Universidad de Sevilla.

Maldonado Trujillo, C., & Pérez Yarahuán, G. (2015). Antología sobre evaluación. La construcción de una disciplina. ((comp.), Ed.) México D.F.: Gobierno y políticas públicas. Centro de Investigación y Docencia Económicas: Centro CLEAR para América Latina.

Martinic, S. (1996). "Evaluación de Proyectos. Conceptos y Herramientas para el Aprendizaje".

May, P. (1993). Claves para diseñar opciones de política. En: L.A. Aguilar (Ed.) Problemas Públicas y Agenda de Gobierno. . Mexico: MAPorrúa Editores.

Moore, M. (1993). Anatomía del problema de la heroína: un ejercicio de definición de problemas. En L. Aguilar, Problemas Públicos y Agenda de Gobierno.

Pastor Albaladejo, G. (2013). El sistema institucional de participación ciudaana: Comunidad de Madrid. Barataria. Revista Castellano-Manchega de Ciencias Sociales(15), 195-205.

Real Academia Española (RAE). (2018). Diccionario de la Lengua Española. (23ª edición. Versión electrónica). Recuperado el 24 de junio de 2019, de https://dle.rae.es/?w=implementaci%C3%B3n

Rein, M., & Schon, D. (2016). Un discurso de políticas que refleja su marco. En La evaluación de políticas. Fundamentos conceptuales y analíticos (págs. 111-146). Corporación Andina de Fomento (CAF) Serie Estado, Gestión Pública y Desarollo en América Latina.

Retolaza, Í. (2018). Teoría del cambio. Una brújula para orientarte en el camino. PNUD/Hivos.

Rossi, P. H., Lipsey, M. W., & Freeman, H. (2003). Evaluation: A Systematic Approach. SAGE Publications, Inc.

Ruano, J. M. (2010). Contra la participación: discurso y realidad de las experiencias de participación ciudadana. Política y Sociedad, 47(3), 93-108.

Ruiz Martínez, A. M. (julio-diciembre de 2013). La participación ciudadana como criterio de valor en evaluaciones de carácter integral. Una aproximación desde la experiencia de AEVAL. Gestión y análisis de políticas públicas, Nueva época(14).

SIEMPRO, UNESCO. (1999). Gestión Integral de Programas sociales orientada a resultados: Manual metodológico para la planificación y evaluación de programas Sociales . México DF.

Silva Jaramillo, S. (2016). Identificando a los protagonistas: el mapeo de actores como herramienta para el diseño y análisis de políticas públicas. Gobernar: The Journal of Latin American Public Policy and Governance, 1(1- Article 7).

Subirats, J. (2001). Nuevos mecanismos participativos y democracia: promesas y amenazas. En J. (. Font, Ciudadanos y decisiones públicas (págs. 33-42). Barcelona: Ariel.

Tamayo, M. (1997). El análisis de las políticas públicas. En R. Bañón, E. Carrillo, & (comps.) (Ed.), La Nueva Administración Pública. Alianza Universidad Textos.

Taylor, S. J., & Bogdan, R. (2008). La entrevista en profundidad. En C. d. docencia, Métodos cuantitativos aplicados. Antología. Chihuahua.

Velasco, M. (18-20 de septiembre de 2007). Distintos instrumentos para un mismo fin. Los instrumentos de las políticas públicas como herramienta para el análisis. Valencia: VII Congreso Español de Ciencia Política y de la Administración (No publicado).

Weiss, C. H. (2016). Preparando el terreno. En C. Maldonado, & G. Pérez, Antología sobre evaluación. La construcción de una disciplina (págs. 27-50). CIDE. Colección Gobierno y políticas públicas.

Wildavsky, A. (1979). The art and craft of policy analysis. London: London and Basingstoke: Macmillan Press.